HONEY BEE (*APIS MELLIFERA* L.) PCR-RFLP DNA MARKERS OBTAINED AND USED FOR THE ANALYSIS OF GENETIC DIVERSITY AND GENE FLOW AMONG NEW WORLD AFRICAN AND EUROPEAN POPULATIONS

By

ALONSO SUAZO CALIX

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2000

This work is dedicated to my wife, Yasmin Cardoza.

ACKNOWLEDGMENTS

I am grateful to my major professor, Dr. Glenn Hall, for his help and support.

Thanks go to the members of my supervisory committee for their help and guidance: Drs. James Nation, Gloria Moore and Jane Brockmann.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

HONEY BEE (*APIS MELLIFERA* L.) PCR-RFLP DNA MARKERS OBTAINED AND USED FOR THE ANALYSIS OF GENETIC DIVERSITY AND GENE FLOW AMONG NEW WORLD AFRICAN AND EUROPEAN POPULATIONS.

By

Alonso Suazo-Calix

May 2000

Chairman: Dr. H. Glenn Hall
Major Department: Entomology and Nematology

Nuclear DNA PCR-RFLP markers were sought and used to distinguish African and European honey bee (*Apis mellifera* L.) subspecies. These markers along with previously described markers were used to genetically characterize and to determine levels of hybridization and gene introgression between New World African and European honey bee populations. Markers were sought in RFLP analyses, by using cloned probes in Southern blots, in RFLP analysis of long (>5kbp) and short (<2kbp) DNA regions amplifiable by long and standard PCR, and in amplified fragment length polymorphism (AFLP) analysis. To facilitate the analyses, markers found using Southern blots and long PCR procedures were made analyzable using standard PCR formats by amplifying the regions containing the informative polymorphic sites. Maps for each polymorphic locus are presented. Characterization of markers specific and predominant to African and European honey bee subspecies is described.

New World honey bee populations from Venezuela, Costa Rica, Honduras, Mexico and the United States (including two populations sampled at two different years near the subtropical-temperate boundary in Texas), representing different lengths of time since African bee invasion, and Old World European and African populations were genetically characterized. Levels of hybridization between New World African and European populations were determined.

Significant differences in the genetic structure were found between Old World European and New World African and European populations ($P<0.01$) but not between New World and Old World African populations ($P>0.05$). From a north to south direction, in neotropical populations, an increase in average frequencies of African-specific alleles was seen, whereas frequencies of east European-specific alleles decreased. West European allele frequencies, however, remained constant, suggesting a high level of population admixture between African and west European bees. Low levels of hybridization between African and European bees were found in neotropical populations and most hybrids were of the African-west European type. In the subtropical-temperate region in Texas, an increase in average frequency of African-specific alleles was seen as reported in the neotropics, however, a high level of hybridization between African and European bees was found.

CHAPTER 1
INTRODUCTION

Biogeography and Distribution of Honey Bee Subspecies in the World.

The honey bee, *Apis mellifera* L., is native to the Old World and was originally distributed in Europe, Africa, and the Middle-East. The wide geographical distribution and adaptation to different environments has resulted in the evolution of 24 subspecies, or races, initially characterized and described using multivariate analysis of morphological characters (Ruttner, 1988). Four taxonomic lineages (or branches) have evolved: near-east, tropical sub-Sahara-African, north African-west Mediterranean (and northwest European) and central Mediterranean-southeast European (Ruttner, 1978). The near-east branch includes the subspecies *A.m. anatolica*, *A.m. adami*, *A.m. cypria*, *A.m. syriaca*, *A.m.caucasica,* and *A.m. meda*. The tropical sub-Sahara-African group includes the subspecies *A.m.adansonii*, *A.m. litorea*, *A.m. monticola*, *A.m. lamarckii*, *A.m. capensis*, *A.m. unicolor,* and *A.m. scutellata*. The north African west-Mediteranean group includes the subspecies *A.m. sahariensis* and *A.m. intermissa*, from north Africa, and *A.m. mellifera* and *A.m. iberica* from western and northern Europe. The central Meditteranean-southeast European group includes the subspecies *A.m. sicula*, *A.m. ligustica*, *A.m. carnica*, *A.m. macedonica,* and *A.m. cecropia*. Throughout this dissertation, "African bees" will refer to the subspecies *A.m. scutellata,* "west-European bees" will refer to the subspecies *A.m. mellifera* and *A.m. iberica*, "east-European bees" will refer to the subspecies *A.m. ligustica* and *A.m. carnica*, and "near-eastern bees" will

refer to the subspecies *A.m. caucasica*. Samples from the other African and European subspecies were not available for this study.

Importation of Honey Bee Subspecies and Establishment in the New World.

The first honey bee colonies were brought to the New World by early European settlers during the first half of the seventeenth century and were primarily of the subspecies *A.m. mellifera*, called German brown or north European black bees (Oertel, 1976). Colonies of *A.m. mellifera* were also imported to North America by the English in 1763 and by the Russians in 1809, but it is not known if these colonies survived (Nelson, 1967). In 1855, an attempt to introduce Italian bees (*A.m. ligustica*) to North America was made, however, not until 1860 was the first successful introduction of this race in North America was reported (Pellet, 1938). Other honey bee subspecies from Hungary and Tunesia (near-east group) and from north Africa followed but with little success, probably due to the poor adaptability of these bees to temperate climates and, consequently, their low performance as honey producers (Pellet, 1938). Egyptian (*A.m. lamarckii*), Cyprian (*A.m. cypria*) and Caucasian (*A.m. caucasica*) bees were also introduced between 1865 and 1905. Of these, only the Caucasian bees were used to any extent by beekeepers. Recent mitochondrial DNA evidence has been reported indicating that honey bees from the African subspecies group, specifically *A.m.lamarckii*, had become established at a low frequency (1.0%) in North America (Schiff and Sheppard, 1993). In North America, the most popular subspecies among beekeepers has become *A.m. ligustica*, Italian bees, followed by *A.m.carnica,* Carniolan bees, which are gentler

and better honey producers than *A.m. mellifera*. In temperate regions of North America, European bees established a substantial feral population composed primarily of *A.m.mellifera, A.m. ligustica,* and *A.m.carnica*.

In the neotropics, the first honey bees, *A.m. mellifera* and possibly *A.m.iberica*, were introduced into Cuba from Spain in 1763 and, later, into Central America. *A.m. mellifera* was also brought to Brazil in 1839 followed by *A.m. ligustica* in 1922 (Kent, 1988). *A.m. iberica* was also brought by the Portuguese to Brazil during the same period (Michener *et al*., 1972). Unlike in North America, *A.m. mellifera* remained the predominant race among beekeepers in the neotropics (Kent, 1988). European bees were poorly adapted to the tropics and did not established self-sustaining feral populations.

African bees, *A.m. scutellata,* were imported to Brazil in 1956 with the intention of interbreeding with and improving the productivity of the European bees. In 1957, swarms from 26 African colonies escaped from apiaries (Kerr, 1967), which resulted in the dramatic spread of African bees in the New World. The African bee migration has been an impressive biological phenomenon in which one subspecies with greater ecological advantages has taken over an existent, albeit largely managed, population of other subspecies. Resident European bees were quickly replaced by Africanized progeny as a result of paternal introgression, that is, African drones mating with European queens (Michener, 1975), and by nest usurpation or invasion of managed European colonies by African swarms (Danka *et al*., 1992; Vergara and Perez de Leon, 1993). As the African bees expanded their range, managed colonies expressed, over time, a significant reduction of morphometric, behavioral, and genetic characteristics typical of European

bees and an increase in African traits. (Quezada-Euan and Medina, 1998; Quezada-Euan and Paxton, 1999; Michener, 1975).

African bees first entered the United States in October 1990 in Hidalgo, Texas (Hunter *et al.*, 1993). The northernmost limit of the African honey bee spread in the United States has been predicted on the basis of a 10°C isotherm line, corresponding to the average high temperature in mid winter, based on their actual distribution in South America (Taylor 1977; Taylor and Spivak 1984). They had been predicted to colonize most of the southern states of the United States by 1995, but, currently, they are only in California, Arizona, Nevada, New Mexico, and Texas. Factors such as the high density of European colonies in Texas and the effect of the parasitic mite *Varroa jacobsonii* are believed to be responsible for the unexpected stall of the expanding African front toward the southeastern states (Sanford, 1994; Eischen, 1999). African bee movement, however, has recently resumed an easterly direction, apparently due to the decimation of feral European colonies by the Varroa mite in Texas (Eischen, 1999). African swarms have been detected near Florida ports from accidental introductions from ships, but it is not known if such introductions will result in the establishment of a substantial feral African bee population (Sanford, 1999). In Puerto Rico, however, a self-sustaining African feral population exists that was established as a result of swarms arriving on ships.

Characteristics of African and European Honey Bees.

African bees evolved in tropical regions, whereas European bees evolved in temperate regions. Unpredictable rains, as well as pollen and nectar availability, have

played an important role in the evolution of African honey bee subspecies. Most of the subspecies variation, based on morphological features, is found in Africa and has been attributed to the diversity of ecosystems and the variability of rainfall (Ruttner, 1998; Rinderer and Hellmich, 1988).

In contrast to African bees, European bees are adapted to more predictable climatic regimes, evolving so that their biological cycles of brood rearing and food storage coincide with temperate seasonal changes (Rinderer and Hellmich, 1988). In general, European bees have better overwintering capabilities, lower swarming and absconding rates and are less defensive than African bees (Michener *et al.*, 1972). The factors responsible for good overwintering capabilities of the European bees include the large size of the individual bees, effective thermoregulation of the brood nest and economical storage and consumption of honey (Winston *et al.*, 1983).

Swarming (natural division of a colony) by African bees is five times more frequent than by European bees. Absconding (abandoning of a nest by a colony) occurs infrequently by European subspecies but commonly by African subspecies, probably due to a higher predation rate of African bees in the tropics (Winston *et al.*, 1983). Other differences between African and European honey bees include the following: developmental period (18 days for African workers compared to 21 days for European workers; Harbo *et al.*, 1981), adult lifespan (shorter in African bees; Winston *et al.*, 1983), selection and size of nesting sites (African bees are usually less selective and usually nest in smaller cavities; Rinderer *et al.*, 1982), defensive behavior (African bees are more defensive; Stort 1974; Michener *et al.*, 1972 and Michener, 1975), size of the

swarms (African bees have smaller swarms; Rinderer *et al.*, 1982) and collection of nectar (African bees collect nectar of lower sugar concentration; Rinderer *et al.*, 1982).

High swarming, absconding, and defensive behavior make African bees more difficult to manage for commercial beekeeping operations (Michener, 1975). In the southern United States, they have already been responsible for several human and domestic animal fatalities. In Florida, they are expected to have a negative impact on the tourism and beekeeping industries. Stinging incidents and even deaths related to African bee attacks will increase, along with liable lawsuits that result from these attacks. Movement of bees from Florida to more northern states, a common practice involving thousand of hives, will be restricted, resulting in economic losses for operations that largely depend on migratory beekeeping.

Methods Used to Identify African and European Honey Bees.

African and European honey bees have been discriminated based on behavioral (Collins *et al.*, 1982), morphometric (Daly and Balling, 1978; Daly *et al.*, 1982), biochemical (cuticular hydrocarbons; Carlson and Bolten, 1983; Lavine *et al.*, 1988; Smith, 1990) and genetic characteristics (allozymes and DNA; Hall, 1986; Hall, 1992b; Hall and Smith, 1991; Sheppard and Berlocher, 1985; Smith and Brown, 1988; Sylvester, 1982). Accurate identification is important for regulatory purposes, certification of honey bee stocks, and research (Page and Erickson, 1985). Reliable identification will be needed to monitor the dynamics of the expanding African population in the United States,

for example, to determine the extent to which African bee characteristics introgress north of the expected hybrid zone in North America.

One of the most notorious behavioral characteristics that distinguish African and European bees is their high defensiveness (Michener, 1975). Defensiveness has been evaluated using quantitative measures of variables such as the time at which the first sting occurred, time taken for the colony to become "very defensive" (African bees have faster reaction time), number of stings (African bees sting in greater numbers), and distance that the bees follow an intruder (Collins and Kubasek, 1982; Collins *et al.*, 1982; Stort, 1974). Another characteristic used for identification is the comb size, which is usually smaller in African bees (Spivak, 1988). Although behavioral characteristics are commonly used, they are subject to variation from environmental factors and management techniques, and they cannot be reliably used to distinguish hybrids.

Morphometric analysis has also been used to discriminate honey bee groups. It measures both genetically based and environmentally induced variation using one (univariate) or several variables (multivariate analysis) (Daly, 1988). Multivariate analysis uses the measurement of different morphological characters to obtain a probability value for a particular sample being African or European. Such analysis has been employed in the Fast Africanized Bee Identification System (FABIS) (Daly and Balling, 1978; Rinderer *et al.*, 1987b). Multivariate analysis requires sophisticated equipment and trained personnel. Multivariate analysis assigns a probability value of being African or European (Daly *et al.*, 1982). Errors can result from the effect of environmental factors on the variation of measurements of morphological characters. As with behavioral methods, morphometrics cannot be used effectively to identify hybrids

and, therefore, is not effective in studying gene introgression and hybridization. Nevertheless, morphometrics has been successfully used to follow changes accompanying the Africanization of some populations in the New World (Boreham and Roubik, 1987). Morphometric values have suggested a European component in the neotropical African population (Quezada-Euan and Hinsull, 1995).

One biochemical method is based on the composition of the insect cuticle (Lockey, 1979). Hydrocarbons are extracted by immersing the insect body in a solvent such as hexane, and the extracted compounds are analyzed using gas chromatography (GC). Differences in the hydrocarbon composition of African and European bees have been found (Carlson and Bolten, 1983; Smith, 1990). The cuticular hydrocarbon composition, however, is also affected by environmental conditions. Bees of the same subspecies that feed on a particular diet can have a different hydrocarbon composition from bees that feed on a different diet. Contamination from pollen, propolis, and plant extracts can obscure the results. Furthermore, the inheritance pattern of the compounds is not known.

Genetic differences between African and European honey bees have been found in the form of protein markers. Protein polymorphism, expressed as allozymes, has been widely used because it is detected by relatively inexpensive and fast procedures. Allozymes are inherited in a clear Mendelian fashion (Del Lama *et al.*, 1990). Allozyme identity in individuals is independent of environmental influences, being the direct product of structural genes (Daly, 1988). However the existence of protein variants is subject to selection, and their detection depends on their expression (Hall, 1988). Honey bees and Hymenopteran insects in general show low levels of allozyme polymorphism

(Crozier, 1970; Metcalf *et al.*, 1975; Sheppard and Berlocher, 1985; Snyder, 1974). This low variability in social Hymenoptera may be due to the exposure of deleterious alleles in haploid males and the relatively small size of the reproductive populations (Graur, 1985; Metcalf *et al.*, 1975). Some of the polymorphic proteins currently available for honey bees include the enzymes esterase, hexokinase, alcohol dehydrogenase, malate dehydrogenase, and a non enzymatic protein P-3 (Contel *et al.*, 1977; Mestriner, 1969; Mestriner and Contel, 1972; Sheppard and Berlocher, 1985).

Unlike protein polymorphism, DNA variation is not necessarily subject to environmental selective pressures, and its detection does not depend on gene expression. Therefore, DNA analysis is the most direct method to establish identities and relatedness among organisms (Hall, 1988). DNA analyses include both mitochondrial (mtDNA) and nuclear (nDNA) DNA. In multicellular animals, mtDNA is maternally inherited (without recombination): the entire mtDNA is transmitted from generation to generation. Changes in mtDNA sequences reflect accumulated mutations (Brown, 1985). Because mtDNA is uniparentally inherited, it cannot be used to study paternal gene introgression and hybridization. Another disadvantage of mtDNA is the relatively small size of the molecule, which limits the number of polymorphisms that can be found. Nevertheless, mitochondrial DNA restriction fragment length polymorphisms (RFLPs) have been found that distinguish African and European honey bees (Hall and Smith, 1991; Smith and Brown, 1988). These markers have been used to test neotropical feral honey bee swarms, thereby demonstrating that the spread of the African bees in the New World has been through unbroken African maternal lineages (Hall and Muralidharan, 1989; Hall and Smith, 1991; Smith *et al.*, 1989).

Compared to mtDNA, nDNA is vastly more abundant in multicellular organisms, allowing for the screening of a practically unlimited number of polymorphisms. Nuclear DNA markers, being maternally and paternally inherited, enable hybridization and gene introgression studies. Nuclear DNA RFLP markers are codominant, thus, allowing the distinction between homozygous and heterozygous individuals. With anonymous cloned probes in conventional Southern blot procedures (Southern, 1975), nDNA RFLP markers that distinguish African and European bees were found and used in the African bee population studies described above (Hall, 1986; Hall, 1990; Hall, 1992b). Southern blot procedures, however, are laborious, expensive, require considerable technical expertise and, thus, limit the application of RFLP markers for population genetic studies in which large sample sizes are required.

Other nuclear DNA markers commonly used to distinguish honey bee subspecies include random amplified polymorphic DNA (RAPD) (Williams *et al.*, 1990). The simplicity of the RAPD analysis makes these markers popular for population genetic studies. However, RAPD markers (Williams *et al.*, 1990) are generally dominant, that is homozygous and heterozygous individuals cannot be distinguished. Other problems associated with the reliability and reproducibility of these markers further limit their application (Black, 1993; MacPherson *et al.*, 1993; Schierwater and Ender, 1993; Weeden, 1992). RAPD markers have been successfully used to map the honey bee genome (Hunt and Page, 1992; Hunt and Page, 1995). RAPD markers that distinguish African and European bees have been reported (Suazo *et al.*, 1998).

Microsatellite markers from single and multiple loci specific to east European, west European, and African bees have been used to characterize Old World European and

African honey bee populations (Estoup *et al.*, 1993; Estoup *et al.*, 1995). Microsatellite analyses are technically more difficult, but simple banding patterns, easy to interpret are produced. However, some microsatellite DNA has been found to be evolutionary unstable (Tautz, 1989).

In past studies, the technical difficulties of RFLP analysis through Southern blot procedures limited the number and specificity of markers that could be found and limited the number of samples that could be tested. Most seriously, at the time of the first studies, no markers specific to west European (*A.m.mellifera* or *A.m.iberica*) had been found (Hall, 1990). The predominant European race in Brazil at the time of African bee introduction was *A.m.mellifera*, but any contribution of this race to the neotropical African population could not have been detected. Subsequently, in Southern blots, RFLP alleles were found at a highly polymorphic locus characteristic of west European bees (McMichael and Hall, 1996). Because the type of analysis limited the number of samples that could be tested and because very complex band patterns characterized the alleles, accurate quantification of allele frequencies was precluded.

To make RFLP markers more productive to use, the polymerase chain reaction (PCR) (Saiki *et al.*, 1988) has been employed. Small DNA regions are amplified and then digested with restriction endonucleases to reveal polymorphisms. This approach is called PCR-RFLP. Compared to Southern blot procedures, PCR-RFLP requires less DNA and no radioactive labeling. It is relatively fast and inexpensive. Large sample sizes can be processed in less time. Thus, PCR-RFLP can greatly facilitate population genetic studies. For honey bees, PCR-RFLP was first used to make their mtDNA easier to analyze (Hall and Smith, 1991; Crozier *et al.*, 1991). More recently, honey bee nuclear

DNA RFLPs, initially found with cloned probes, have been made PCR analyzable: one locus with African-specific markers (Hall, 1998) and three loci with east European markers (Hall, in preparation).

Processes Involved in the Spread of the African Bees in the New World.

In the forty three years since the African bees were first released, they have successfully established self-sustaining feral populations in the tropical and subtropical regions of the New World. The processes involved in the spread and establishment of the African bees are not well understood. There are two controversial views (reviewed in Hall, 1992a). One view proposes that the primary driving force responsible for the spread has been through African drones mating with European queens, resulting in a population called a "hybrid swarm" (Rinderer *et al.,* 1985). These hybrid colonies, which were originally European and have been called "Africanized bees", further perpetuated the spread of African genes. According to this view, differential drone production and social parasitism are the important factors that have contributed to the Africanization process (Rinderer *et al.*, 1985; Rinderer *et al*., 1986; Rinderer *et al.,* 1987a). The other view is that the primary driving force has been African maternal expansion through feral swarms. This view was based on strong mitochondrial (mtDNA) evidence from independent studies (Hall and Muralidharan, 1989; Hall and Smith, 1991; Smith *et al*., 1989). Virtually all mtDNA in feral swarms in Brazil, Venezuela, and Mexico were of the African type. This view also holds that the neotropical population is not a hybrid swarm. Hall (1990) used both mtDNA and east European nDNA markers to test a group

of managed European colonies in southern Mexico established before, but sampled after, African bees entered the region. The results showed that the colonies retained European mtDNA but had low levels of European nDNA markers, probably the result of European queens mating with African drones. Feral African swarms were also tested. Swarms near the expanding front had low levels of European markers, whereas those sampled farther behind the front had virtually no European markers. These findings pointed to asymmetric gene flow: substantial African paternal introgression into managed European colonies but limited paternal introgression from the European colonies into the feral African bee population (Hall, 1990). Similar findings were obtained by others with allozyme markers (Taylor *et al*., 1991).

Selection against European bees and their hybrids in tropical conditions probably contributed to the rapid Africanization of the European population. However, low hybrid survivability may also be a result of low metabolic capacity compared to either African or European bees (Harrison and Hall, 1993). Other factors, such as queen developmental time (DeGrandi-Hoffman and Watkins, 1998; DeGrandi-Hoffman *et al*., 1998), may have also contributed to the Africanization process. A honey bee queen mates with an average of a dozen drones (Adams *et al*., 1977). Among daughter queens from mothers mated to both African and European drones, those with African paternity will develop faster and will emerge first, killing their slowly developing sisters still in their cells. Through kin selection, workers with African paternity may preferentially, and more successfully, rear closely related sister queens. There is yet no evidence to support this possibility. Although European genes would be expressed in the worker population, these queen selective processes would eliminate them from the reproductive population.

In the subtropical-temperate boundary region, where both African and European bees will be equally adapted, a hybrid zone is expected be formed (Taylor and Spivak, 1984). Bees with European genotypes will be better adapted to the temperate climate north of this hybrid zone, whereas those with African genotypes will be better adapted to the subtropical conditions south of the zone.

Needs for Additional Studies

More DNA markers that can be easily used have been needed to distinguish African and European bees and their hybrids. In addition to research, DNA markers are needed for regulatory identification of African bees in the U.S.A and for stock certification.

Additional studies have been needed to obtain a more accurate measure of subspecies contributions, introgression patterns, and the resulting genetic structure of African and European honey bee populations in the New World. The subspecies contribution of the feral population in the neotropics needs to be better documented. Also the extent that the African European genotypic composition affects the fitness of hybrid colonies remains to be determined. Hybrid fitness in either tropical or temperate conditions must be established before implementing breeding programs proposed to produce better adapted and gentler hybrids for commercial beekeeping. Many hybrid studies will depend upon genetic markers that are specific to African and European honey bee subspecies, that involves cost effective and accurate methods to screen honey bee populations.

Goals of This Research.

My doctoral research had two major goals. The first goal was to employ PCR to augment the collection of markers. The second goal was to apply the markers in an advanced population genetic study involving a broader sampling of New World bees. In accomplishing the first major goal, nuclear DNA PCR-RFLP markers were obtained with three different approaches. First, markers found using conventional Southern blot procedures were made PCR analyzable by isolating and amplifying the regions that contained most of the polymorphisms (Chapter 2). Second, new markers were sought and found in DNA initially amplified using standard and long-PCR rather than in DNA detected with cloned probes using Southern blot procedures (Chapters 3 and 4). The effectiveness of using long versus standard PCR protocols was also compared (Chapter 3). To reduce more stringent requirements used in long-PCR protocols, markers found with long-PCR were made analyzable by standard PCR protocols (Chapter 5). Finally, markers were also obtained after modifying the amplified fragment length polyorphism (AFLP) procedure, a technique that can detect a large amount of polymorphism in plant and animals (Chapter 6). To accomplish the second main goal, the PCR-RFLP markers found in this study were used together with previously described PCR-RFLP markers to test samples from Venezuela, Costa Rica, Honduras, Mexico, and the U.S.A, as well as Old World African samples of *A.m.scutellata* and Old World European samples of *A.m.mellifera*, *A.m.iberica*, *A.m.ligustica*, *A.m.carnica,* and *A.m.caucasica*. Samples from two populations were included in the analysis, taken at two different periods near the subtropical-temperate boundary in Texas.

As an outcome of this study, much more accurate estimates of gene frequencies and levels of hybridization were determined. New important information was obtained that either confirmed or required modifications of previous conclusions.

CHAPTER 2
CHARACTERIZATION OF POLYMORPHIC PROBE P227

Introduction

Conventional methods used to search for RFLP markers make use of Southern blot procedures (Southern, 1975) which typically use cloned DNA fragments are used as probes. Restriction fragment length polymorphisms (RFLPs) in mitochondrial (mtDNA) and nuclear DNA (nDNA) have been found that distinguish African and European honey bee subspecies (Hall 1986; Hall 1990; Hall 1992b; Hall 1998; Hall and Smith 1991). These markers have been valuable in revealing processes involved in the African bee spread in the New World and in detecting accidental introductions of African bee swarms in Florida, USA (Hall 1990, Hall 1992a; Hall and Muralidharan 1989; Sanford, 1999; Smith *et al*., 1989). Southern blot analysis, however, is laborious, time consuming, expensive and requires considerable technical expertise. Large amounts of DNA are also required, thus, the number of enzymes that can be screened to reveal polymorphisms in a particular sample is limited to the amount of DNA that can be extracted. Therefore, these analyses are difficult to use especially when a large number of samples must be tested, as in population genetic studies. With the implementation of the polymerase chain reaction (PCR) (Saiki *et al*., 1988), samples can now be amplified in a relatively short period of time with little amounts of DNA and at a relatively low cost. Large number of samples can be tested more effectively. Thus, PCR analyzable RFLP markers are more suitable for population genetic studies.

In this chapter, a highly polymorphic locus, P227, detected using conventional Southern blot procedures, is characterized, and its informative polymorphic section made analyzable by PCR. This enabled the section to be screened with additional enzymes than those used in the initial screening using Southern blots. Locus P227 was found to have alleles specific to African, east and west European honey bee subspecies seen with the restriction endonucleases *Alu*I and *Hinf*I. Frequency data are reported for both African and European alleles detected with *Alu*I and *Hinf*I in African and European populations.

Materials and Methods

Honey Bee DNA Samples and DNA Isolation

Samples used were drones (haploid males) and workers (diploid females) from four different locations in the Transvaal of South Africa: White River, Louis Trichardt, Warm Baths and Pretoria (collected by H. G. Hall, University of Florida, Gainesville, FL), three locations from the United States (U.S.A): Florida, Kansas (provided by O. R. Taylor, University of Kansas, Lawrence, KS) and from a closed breeding population representing bee stocks from across the United States in Tucson, Arizona (Page *et al.*, 1982; Severson *et al.*, 1986) (provided by G. Loper, G. Waller, J. Martin and E. Erickson, Carl Hayden Bee Research Center, USDA-ARS). Adult worker samples from France (*Apis mellifera mellifera*), Italy (*A.m. ligustica*), Spain (*A.m. iberica*) and Austria (*A.m. carnica*) (provided by D. R. Smith, University of Kansas, Lawrence, KS) and from a

collection of bees collected in Russia and maintained in France (*A. m. caucasica*) were also used to determine the specificity of the European markers. One to three workers were used per colony from a total of 67 colonies from South Africa, 44 colonies from the USA, nine colonies from France, 41 colonies from Italy, 26 colonies from Austria, 15 colonies from Spain and 7 colonies from France (from a collection). European samples were grouped as either west European (*A.m. mellifera* and *A.m. iberica*) or east European (*A.m. ligustica, A.m. carnica*). *A.m. caucasica* was grouped as a fourth near-east lineage as suggested by morphometry and more recently, by mitochondrial and nuclear DNA markers (Arias and Sheppard, 1996; Ruttner, 1988). DNA was isolated from drone pupae and from thoracic muscles of adult honey bee workers as described in Hall, (1986) and Hall, (1990), and by using anionic columns as suggested by the manufacturer (QIAGEN, Palo Alto, CA). Drone samples (as pupae) were primarily used in Southern blot analysis because of their simple haploid pattern of restriction fragments.

Probe Description and Screening for Polymorphisms.

Construction of a plasmid library of cloned honey bee fragments, electrophoresis of digested DNA, blotting, probe labeling, prehybridization, hybridization and autoradiography were as described in Hall 1986 and Hall 1995. A honey bee genomic DNA library was generated from *Pst*I digested honey bee DNA cloned into *Pst*I digested pBR322 as plasmid vector. The bee DNA fragments were from low-copy number sequences and were larger than 4 kilobases (kb) (Hall, 1986). Probe 227 (P227) selected

from this library had a 9.2 kilobase (kb) insert. To screen for enzymes that reveal polymorphisms, aliquots of 5μg of DNA from one African and from one European drone sample was separately digested with nine different restriction endonucleases: *Msp*I, *Sau*96I, *Alu*I, *Hae*III, *Hha*I, *Nci*I, *Mbo*I, *Hinf*I and *Dde*I. The DNA digested with the different enzymes was electrophoresed in separate wells of a 2% agarose gel and transferred to a nitrocellulose membrane (blots) (Genescreen plus, NEN Research products, Boston, MA). Probe was labeled by Nick translation as described (Sambrook *et al*., 1989). Blots were prehybridized, hybridized with radioactively labeled P227 as probe and exposed to autoradiographic film (Kodak X-Omat, Rochester, NY). From this initial blot, enzymes that revealed distinctive polymorphisms between the African and European honey bee DNA samples were selected for further testing. Additional blots (three to four) containing DNA from nine African and nine European samples digested with the enzymes that revealed polymorphisms during the initial screening were used to determine allele frequencies.

Subcloning

In this chapter and in Chapter 5, the term "region" will refer to the entire locus under study. Alleles will be the different banding patterns found at a particular locus or "region". The conversion of markers found with Southern blot procedures and long PCR required the cloning of "sections" of each region. Alleles resulting from RFLP analysis of these sections will be referred to as "sub-alleles". Thus, the terms "regions" and "alleles" are analogous to "sections" and "sub-alleles".

Restriction sites of the enzymes contained in the polycloning site of the plasmid vector pGem3Z (Promega corporation, Madison, WI) were mapped in P227. To subdivide the entire probe into several sections of sizes amplifiable by standard PCR protocols, three restriction endonucleases were selected, based on the size of fragments they generated and their relative position in P227: *Xho*I, *Sph*I and *Hind*III. The relative positions of the selected enzymes were mapped in P227 using standard double digestion procedures (Sambrook *et al*., 1989). These enzymes divided P227 in four different sections: section 1 (S1), from the terminal *Pst*I site (position 0) to the *Xho*I site (position 2500)(2.5kb), section 2 (S2) from the *Xho*I site to the *Sph*I site (position 3900)(1.4kb), section 3 (S3) from the *Sph*I site to the *Hind*III site (position 7100)(3.2kb) and section 4 (S4) from the *Hind*III site to the second terminal *Pst*I site (position 9200)(2.1kb) (Figure 2-1). Each section was subcloned into pGem3Z vector (Promega corporation, Madison, WI) and transformed into *E. coli* competent cell DH5α (Lifetechnologies, GIBCO BRL, Gaithersburg, MD). To clone the sections, 500ng of P227 and 500ng of pGem3Z (Promega corporation, Madison, WI) were digested with the enzymes that flanked the corresponding section. The digestion was done in 50μl of total volume with 10 units of each enzyme and using the buffers and conditions following manufacturers recommendations. Fragments for both the insert (section of P227 to subclone) and the vector (pGem3Z) were separated in 1% agarose gels, excised with a blade, and placed in dialysis bags. Bags were filled with 1X concentrated Tris Borate Buffer (TBE). To elute the DNA fragments from the agarose gel slice, the dialysis bag was placed in an electrophoresis unit and electrophoresed at 4V/cm for 2hrs. The buffer containing the eluted DNA fragments was collected, and DNA fragments were purified using anionic

resin columns (DE52, Whatman, Maidstone, England) as described (Sambrook *et al.*, 1989). DNA fragments were precipitated with 500µl of 95% ethanol, washed with 500µl of 75% ethanol and resuspended in 20µl of water. Equal amounts of both purified vector and insert (1ng of each) were ligated in a total of 20µl with 1unit of T4 ligase (Boehringer Mannheim, Indianapolis, IN) and buffer conditions as suggested by the manufacturers. Ligation reactions were performed at 15°C for 16hrs. Colony transformation was done with 100µl of competent cells DH5α (Lifetechnologies, Gibco BRL, Gaithersburg, MD). To transform the cells, 4µl of the ligation mixture were place on top of the competent cells and mixed gently. Cells were incubated on ice for 30 min, heat shocked at 42°C for 45 sec and placed on ice for 2 min. Following incubation on ice, 900µl of SOC medium (Sambrook *et al.*, 1989) were added, and the mix was incubated at 37°C for 1hr. The cells were plated on SOB medium (Sambrook *et al.*, 1989) and allowed to grow overnight. Cells were then screened for white or blue colonies for the presence or absence of inserts, respectively. Subclones corresponding to each section were used separately as radioactive probes in the original blots, to find the sections that contained the informative polymorphic sites. Subclones revealing informative polymorphisms were sequenced about 500 bases from their terminal ends (University of Florida, Interdisciplinary center for biotechnology research ICBR, Gainesville, FL). From these terminal sequences, primer sequences were selected using the computer software package OLIGO 5.0 (National Bioscience, Plymouth, MN).

Amplification and RFLP Analysis of PCR Products

Almost the entire S3 of P227 with the adjacent end of S4 was amplified as one segment, referred to as Section 3 extended (P227-S3xt). Primers used were [1] 5'-AGAAGGAAAGAAGAAACGGATGAAC-3' within S3 (position 4030) near the *Sph*I site and, [2] 5'-CGGAGGAGTGGTAATAATGGAAGC-3' within S4 (position 7496) near the *Hind*III site (Figure 2-1). Amplifications were performed in 25µl of 50mM Tris-HCl pH 9.2, 16mM $(NH_4)_2SO_4$ and 1.8mM $MgCl_2$ containing 200µM of each dNTP's, 250nM of each primer with 1.12 units of *Taq* DNA polymerase and 0.12 units of *Pwo*I DNA polymerase taken from a stock mixture containing nine units of *Taq* DNA polymerase and one unit of *Pwo*I DNA polymerase (Boheringer Mannheim, Indianapolis, IN). Fifty to 250ng of genomic DNA was used per reaction (the highest amount of DNA was used when DNA was extracted from thoracic muscles of adult honey bees). Amplifications were performed in a PTC-100 or PTC-200 thermocycler (MJ Research, Watertown, MA) and the following profile was used: 95°C for 2 min followed by 33 cycles each at 94°C for 45s, 57°C for 45s and 68°C for 4min with a final extension at 68°C for 10min. A second section near the *Sph*I terminus of S3 (position 4030 to 4300) referred to as P227 section 3 terminal (P227-S3ter) was amplified using primer [1] from P227-S3xt and primer [3] 5'-CACTTTTTCAAAAGAGCGTGCAA-3' (position 4300) (Fig. 2-1). The reaction solution to amplify P227-S3ter was as with P227-S3xt, except that 1.0mM $MgCl_2$ and one unit of *Taq* DNA polymerase were used (Lifetechnologies GIBCO BRL, Gaithersburg, MD). The amplification profile for P227-S3ter was as follows: 95°C for 2 min followed by 35 cycles of 94°C for 45sec, 60°C for 45sec and 72°C for 30sec with a final extension of 5 min at 72°C. P227-S3xt was screened for 22

restriction endonucleases, some of which had been used in the initial screening using southern blots: *Hae*III, *Hinf*I, *Msp*I, *Dra*I, *Ava*II, *Bam*HI, *Eco*RI, *Dde*I, *Afl*II, *Bgl*I, *NrU*I, *Stu*I, *Bcl*I, *Nhe*I, *Bst*XI, *Eco*RV, *Spe*I, *Hpa*I, *Nsi*I, *Bgl*II, *Nco*I and *Sty*I. Five µl of amplified products were digested with 5µl of a digestion mixture containing 1 unit of enzyme, 2mM DTT, 25µg of BSA, and restriction digestion buffer to a final concentration of 1X as suggested by the manufacturers. Digested PCR products were separated in 2% agarose gels, stained with ethidium bromide, and visualized over UV light. The initial screening was done with nine African and nine European samples to increase the chances of finding polymorphisms, compared to one African and one European drone used in the initial screening with Southern blots. Enzymes that revealed polymorphisms in all of one group and not the other were tested in a larger sample size consisting of 48 African and 48 European drone samples. The restriction enzymes *Alu*I and *Hinf*I were found to generate informative polymorphisms.

PCR products were used to map the restriction sites for all the alleles of the enzymes that revealed significant polymorphisms following the method described by Her and Weinshilboum, 1995 and described in more detail in Chapter 5.

Sub-Allele Count and Data Analysis

The different fragment patterns generated by *Alu*I and *Hinf*I were identified in both drones and workers. Each pattern identified is considered a sub-allele. Sub-alleles were named using the enzyme name followed by a letter, for example, *AluI-A*. Drones (haploid parthenogenetic progeny) were used to determine the queen genotypes from which sub-allele frequencies were obtained. Two to five drones per colony were used.

When three or more drones where found to have the same sub-allele, the queen's genotype was determined as homozygous for that sub-allele, and the sub-allele was counted twice. If two different sub-alleles were found, each one was scored once. Each drone was assigned a combination of two letters corresponding to the sub-alleles produced by *Alu*I and *Hinf*I. These combinations are referred to as composite sub-alleles. The coupled or repulsion phase of sub-alleles produced by two restriction enzymes, thus the identity of the composite sub-alleles, could not be determined in diploid individuals i.e. workers. For example, a worker with genotype *Alu*I-*A:Alu*I-*B/Hinf*I-*A:Hinf*I-*B* for both *Alu*I (heterozygous for the *Alu*I) and *Hinf*I (heterozygous for *Hinf*I) will have four possible composite sub-alleles: *Alu*I-*A/Hinf*I-*A*, *Alu*I-*A/Hinf*I-*B, Alu*I-*B/Hinf*I-*A* and *Alu*I-*B/Hinf*I-*B*. In this case, a cis configuration (*Alu*I-*A*:*Hinf*I-*A*) cannot be distinguished from a trans configuration (*Alu*I-*A*: *Hinf*I-*B*). Frequency data obtained from composite sub-alleles in drone samples was used to generate a contingency table. Sub-allele association was tested by testing the contingency table for independence using a Fisher's exact test (under the null hypothesis of no association) based on the metropolis algorithm using the computer software package RXC (M.Miller, University of Arizona). Frequency data from workers was obtained from direct counts of *Alu*I and *Hinf*I sub-alleles. Observed and Expected heterozygosities (H_{obs} and H_{exp}) were calculated as described in Nei, 1987 and Nei, 1978 using the computer software package GENEPOP version 1.2 (Raymond and Rousset, 1995).

Results

Initial screening, using P227 as probe in Southern blots with DNA from one African and one European sample digested with 9 enzymes, revealed informative polymorphisms with *Alu*I. Further analysis using blots containing nine African and nine European samples showed a total of nine alleles of which six were African-specific and three were European-specific (Figure 2-2). All polymorphic bands resulting from *Alu*I digests were detected when subclones of both section 3 (S3) and section 4 (S4) of P227 were used as probes. S4 detected only one *Alu*I polymorphic site that was located at the terminal end, adjacent to S3 (determined from sequencing data). RFLP analysis of the amplifiable P227-S3xt also revealed polymorphisms with *Hinf*I.

*Alu*I Polymorphisms

*Alu*I digestions of P227-S3xt region revealed nine sub-alleles (Figure 2-3). Sub-alleles *Alu*I-*A*, *C*, *D*, and *J* were found predominantly in the South African population at a collective frequency of 52.4% (Table 2-1). *Alu*I-*J*, found at a 39.5% frequency in the African samples, was absent from both east and west European types but present at a very low frequency (0.8%) in the U.S.A samples. Two sub-alleles (*Alu*I-*I* and *E*) were predominantly found in the European samples. *Alu*I-I was found in east and west European samples at frequencies of 100.0% and 79.1% in *A.m.ligustica* and *A.m.carnica,* respectively, and 65.0% and 30.7% in *A.m.mellifera* and *A.m.iberica,* respectively. *Alu*I-*I* was absent in bees of the near-east lineage group of *A.m.caucasica*. *Alu*I-*E* was *A.m.caucasica*-specific and was found at a frequency of 92.1% among the *A.m.caucasica*

samples. Sub-alleles *Alu*I-*I* and *E* were found at a collective frequency of 83.5% in the U.S.A samples.

Sub-alleles *Alu*I-*A1*, *D* and *I* have unique fragments of 1000 (*A1*), 2834 (*D*) and 1650 bp (*I*), respectively. Sub-alleles *Alu*I-*B* and *E*, *Alu*I-*E* and *F* and, *Alu*I-*F* and *J* are characterized by fragments of 2574, 340 and 2354bp, respectively. An *Alu*I site, located 220 bases from primer [1] within S3 (close to the *Sph*I site) is present in sub-alleles *Alu*I-*A*, *A1*, *C*, *D*, *F* and *J*, most of which are African-predominant or African-specific (except *Alu*I-*A1* and *F*), and absent in *Alu*I-*E*, *I* (European-predominant sub-alleles) and *B* (common sub-allele). Sub-alleles *Alu*I-*J* (African-specific) and *B* (common) have very similar banding patterns but differ by the presence and absence, respectively, of this restriction site, at position 220, which generates the characteristic fragment of 2574bp in *Alu*I-*B* and 2354bp in *Alu*I-*J*. The 220bp fragment in *Alu*I-*J* comigrates with a separate 225bp fragment, and both appear in agarose gels as only one band, making the distinction between *Alu*I-*B* and *J* difficult. Distinction between these two sub-alleles can be made by amplifying and digesting a shorter section (P227-S3ter). The presence of a 280bp fragment (uncut) identifies the *Alu*I-*B* sub-allele (common sub-allele) whereas its absence identifies the *Alu*I-*J* sub-allele (African-specific) (Figure 2-3).

Genotypes that cannot be distinguished include: *A*/*A* from *A*/*C*, *A1*/*A1* from *A1*/*C* and *A*/*I* from *C*/*I*. The inability to distinguish these genotypes can create a bias in the frequency of both *Alu*I-*A* and *Alu*I-*A1*, however, *Alu*I-*C* was found at very low frequencies in our samples (only in one drone, corresponding to a frequency of 0.3%).

*Hinf*I Polymorphisms

Five sub-alleles were revealed with *Hinf*I, of which three were European-specific (*Hinf*I-*F*, *G* and *H*). *Hinf*I-*F* was only found in east European samples (56.1% in *A.m. ligustica* and 66.7% in *A.m. carnica*) and in 35.9% of the *A.m. caucasica* samples. *Hinf*I-*G* was fixed in the west European samples (100.0% in both *A.m.mellifera* and *A.m.iberica*) and was found at an 18.8% frequency in the *A.m. carnica* and at a 10.3% frequency in the *A.m caucasica* samples. *Hinf*I-*F*, *G* and *H* were found at a collective frequency of 74.2% in the U.S.A samples (Table 2-1). *Hinf*I-*H* was found only in the U.S.A samples at an 11.7% frequency. One African-specific sub-allele was found (*Hinf*I-*C*) at a frequency of less than 1.0% in the South African samples.

Sub-alleles *Hinf*I-*C*, *F*, *G* and *H* were characterized by fragments of 400 (*C*), 380 (*F*), 250 and 225 (*G*), and 470bp (*H*) respectively (Figure 2-4). Undistinguishable genotypes include *C*/*C* from *C*/*D* and *F*/*F* from *F*/*D*.

A strong association between *Alu*I and *Hinf*I sub-alleles was found ($P<0.0001 \pm 0.000$). These sub-allele associations reinforce the specificity and hence, the value of these markers for diagnostic purposes. For example, *Alu*I-*I*, a predominantly east European sub-allele was generally found associated with *Hinf*I-*F*, an east European-specific sub-allele (Table 2-2). African-predominant or specific sub-alleles were found associated with African or common sub-alleles. *Hinf*I-*H* may not be a European sub-allele. It was not found in samples from Europe, and it was associated with an African-specific sub-allele (*Alu*I-*J*).

Observed heterozygosity values for sub-alleles found with *Alu*I were higher in South African and west European bees compared to east European and near-east bees. Compared to the *Alu*I sub-alleles, heterozygosity values found with *Hinf*I sub-alleles was higher in east European bees compared to South African and west European (Table 2-1).

Discussion

We have employed the polymerase chain reaction to amplify a polymorphic region, P227, initially detected by Southern blot analysis, with sub-alleles specific to or predominant in both African and European honey bee subspecies. More importantly, P227 effectively distinguishes bees in the African (*A.m. scutellata*), west European (*A.m. mellifera* and *A.m. iberica*) and east European (*A.m. ligustica, A.m. carnica*), and near-east (*A.m. caucasica*) group of subspecies. Sub-allele frequency data obtained from *Alu*I and *Hinf*I digests of P227-S3xt reflects the genetic composition of bees from the United States and the contributions of different subspecies introduced by European settlers to North America (Nelson, 1967; Oertel, 1976; Pellet, 1938). The combined European sub-allele frequency in the North American population is consistent with that found for alleles at other polymorphic loci (Hall, 1990; McMichael and Hall, 1996). Most of these sub-alleles correspond to east European-specific sub-alleles, which reflects the preference of the Italian honey bee, *A.m.ligustica,* among US beekeepers. West European bees were introduced to North America prior to the introduction of east European bees, but, because of their defensive behavior, they were not as popular among beekeepers (Oertel, 1976; Pellet, 1938). This is shown by the lower frequency of the west European-specific sub-alleles found with *Hinf*I (*Hinf*I-*G*). Low frequencies of the *A.m.caucasica* sub-allele

AluI-E in the U.S.A samples also show a genetic contribution of this subspecies in the U.S.A.

Honey bees from Egypt (*A.m. lamarckii*), and possibly from other African subspecies, were introduced in the U.S.A in the early 18th century (Oertel, 1976; Pellet, 1938). Low adaptability to temperate conditions probably resulted in failure to establish a substantial managed or feral population, as did the west and east European subspecies. However, it is possible that alleles from these African subspecies could have introgressed and persisted at low frequencies in feral and managed colonies in North America. Sub-alleles found at very low frequencies in South Africa and the U.S.A (example: *AluI-F*) could possibly reflect the contribution of these non-*A.m.scutellata* African subspecies in the New World. Introduction of non-*A.m.scutellata* African subspecies to America has been suggested based on mitochondrial DNA evidence in South American populations (Arias and Sheppard, 1996). Migratory beekeeping practices could also contribute to the presence of low frequency sub-alleles in other populations as, for example, the regular movement of *A.m. ligustica* between Italy and France where *A.m. mellifera* is predominantly found (Estoup *et al.*, 1995). African subspecies such as *A.m. intermissa* brought from Morroco (Northwest Africa) into the Iberian peninsula where *A.m. iberica* is predominantly found (Garnery *et al.*, 1978a and b; Franck *et al.*, 1998). European-predominant sub-alleles are also found at low frequencies (less than 4%) in the South African population. European bees were imported to South Africa but did not establish a feral or managed population because of their poor adaptability to tropical conditions (Fletcher, 1973; Fletcher, 1978). From allozyme (Smith and Glenn, 1995), mtDNA (Smith *et al.*, 1991), and microsatellite (Franck *et al.* 1998) studies, African-specific

alleles have been found in west European populations, particularly *A.m. iberica*. Recent studies have proposed that the Iberian peninsula could be an intergradation zone in which west European-specific and African alleles are found, the latter having introgressed from North Africa (Smith *et al*. 1991 and Franck *et al*. 1998). *Alu*I-*A1* is found at a high frequency in *A.m. iberica*, at intermediate frequencies in *A.m.scutellat,a* and at very low frequencies in the U.S.A samples. *Alu*I-*A1* may be a sub-allele that originated from a north-western African subspecies and introgressed north into the Iberian peninsula and south into the *A.m. scutellata* populations. This hypothesis cannot be substantiated with our current data.

African and west European-specific sub-alleles have been more difficult to find than east European-specific sub-alleles. High swarming, high genetic diversity and migratory beekeeping practices may have allowed for a greater blending of more diverse populations, and this may have contributed to the difficulties of finding African-specific markers. Markers found with P227 include not only African-specific but also west and east European-specific alleles, which increases the value of this locus thereby facilitating the use of these markers in studying gene flow or genetic variation between African and European populations in the New World.

This polymorphic region, P227, was successfully detected using Southern blot procedures, and the sections containing the polymorphic sites were made analyzable by PCR. PCR facilitates RFLP analysis of the polymorphic sections by making the procedure faster, easier to use, less expensive and by reducing the quality and quantity of DNA required. PCR analyzable RFLPs of P227 are codominant markers, (homozygous

and heterozygous individuals can be distinguished), thus, allowing these markers to be used for parental analysis and hybridization studies.

Table 2-1. Sub-allele frequencies, observed and expected heterozygosity (H_{obs} and H_{exp} respectively) and, collective sub-allele frequencies in African (C_{AF}), east European (C_{EE}), west European (C_{WE}) and, near-east, *A.m.caucasica* (C_O) in New World honey bee populations for locus P227. "N" corresponds to the number of workers.

Enzyme Allele	Populations						
	African		West European		East European		Other
	South Africa (*A.m.scutellata*)	United States	France (*A.m.mellifera*)	Spain (*A.m.iberica*)	Italy (*A.m.ligustica*)	Austria (*A.m.carnica*)	Russia (*A.m.caucasica*)
*Alu*I							
(N)	127	60	10	13	42	24	19
A[*]	0.1260		0.0500				
A1	0.3307	0.0250	0.3000	0.6923		0.1667	
B	0.0551	0.0583				0.0417	0.0789
C[*†]							
D[*†]							
E[§]		0.1000					0.9211
F	0.0512	0.0667					
I[§]	0.0394	0.7417	0.6500	0.3077	1.0000	0.7917	
J[*]	0.3976	0.0083					
C_{AF} (*A,J*)	0.5236	0.0083	0.0500				
C_{EE} (*I*)	0.0394	0.7417*	0.6500	0.3077	1.0000	0.7917	
Co (*E*)		0.1000					0.9211
H_{obs}	0.4567	0.3000	0.3000	0.4615	0.0000	0.1667	0.1579
H_{exp}	0.7094	0.4314	0.4850	0.4260	0.0000	0.3438	0.1454
*Hinf*I							
(N)	127	60	12	14	41	24	19

Table 2-1 continued.

C							
D	1.0000	0.2583			0.4390	0.1458	1.0000
F[§]		0.4167			0.5610	0.6667	
G[§§]		0.2083	1.0000	1.0000		0.1875	
H		0.1167**					
C_{EE} (*F*)		0.4167			0.5610	0.6667	
C_{WE} (*G*)		0.2083	1.0000	1.0000		0.1875	
H_{obs}	0.0000	0.5000	0.0000	0.0000	0.8780	0.4167	0.0000
H_{exp}	0.0000	0.7026	0.0000	0.0000	0.4926	0.4991	0.0000

¥: African-specific or predominant sub-alleles.

§: East European-specific or predominant sub-allele.

§§:West European-specific or predominant sub-allele.

*: Does not include the frequency of the *A.m.caucasica* sub-allele *Alu*I-E.

**: Found only in the U.S.A.

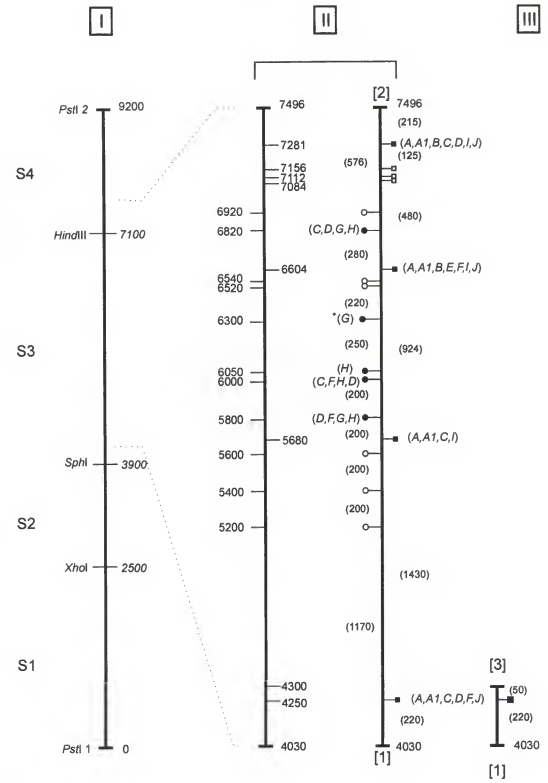
†: Sub-alleles not detected in workers and found in only two drones.

Table 2-2. Contingency table of composite sub-allele frequencies for *Alu*I (columns) and *Hinf*I (rows) sub-alleles. Sub-allele association frequencies were determined in drone samples.

	$A^{¥}$	C^{*}	D^{*}	J^{*}	E^{*}	$I^{§}$	F	$A1$	B
C^{*}				2					
$F^{§}$						28			
$G^{§§}$						4			
H^{**}				2			2		
D	5	1	1	9	5	3	2	15	5

¥ : African-specific or predominant sub-alleles.
§: East European-specific or predominant sub-allele.
§§:West European-specific or predominant sub-allele.
*: *A.m.caucasica* –specific sub-allele.
**: Found only in the U.S.A.
$P<0.0001 \pm 0.0000$ (Fisher's exact test).

Figure 2-1. Restriction enzyme cleavage site map for locus P227 and the amplifiable polymorphic sections. I) General map of locus 227 with the enzymes used to divide and subclone the different sections. II) Position of *Alu*I and *Hinf*I sites in the polymorphic section 3 extended (P227-S3xt). III) Amplified sections of P227-S3ter used to distinguish sub-alleles *Alu*I-*B* and *Alu*I-*J*. Locus P227 was divided in four sections, S1 to S4 defined by the restriction sites *Xho*I, *Sph*I, and *Hind*III. Squares and circles correspond to *Alu*I and *Hinf*I sites, respectively. Open and closed symbols represent monomorphic and polymorphic sites, respectively. Letters in parentheses are the sub-alleles that have the indicated *Alu*I and *Hinf*I sites. Position of restriction sites, in base pairs, are indicated to the right (*Alu*I) or left (*Hinf*I) side of the map. Primers used to amplify each region are indicated in brackets above and below each diagram. A single individual star indicates the polymorphic *Hinf*I site that is west European-predominant (*A.m.mellifera* and *A.m.iberica*). The amplified sections of P227-S3ter (far right) includes the polymorphic site that distinguishes sub-alleles *Alu*I-*B* (common) and *Alu*I-*J* (African-specific).


I
II
III
PstI 2
9200
HindIII
7100
SphI
3900
XhoI
2500
PstI 1
0
S4
S3
S2
S1
7496
7281
7156
7112
7084
6920
6820
6604
6540
6520
6300
6050
6000
5800
5680
5600
5400
5200
4300
4250
4030
[2]
7496
(215)
(A,A1,B,C,D,I,J)
(125)
(576)
(480)
(C,D,G,H)
(280)
(A,A1,B,E,F,I,J)
(220)
*(G)
(250)
(924)
(H)
(C,F,H,D)
(200)
(D,F,G,H)
(200)
(A,A1,C,I)
(200)
(200)
(1430)
(1170)
(A,A1,C,D,F,J)
(220)
4030
[1]
[3]
(50)
(220)
4030
[1]

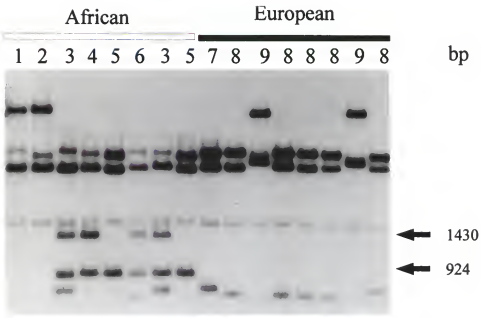
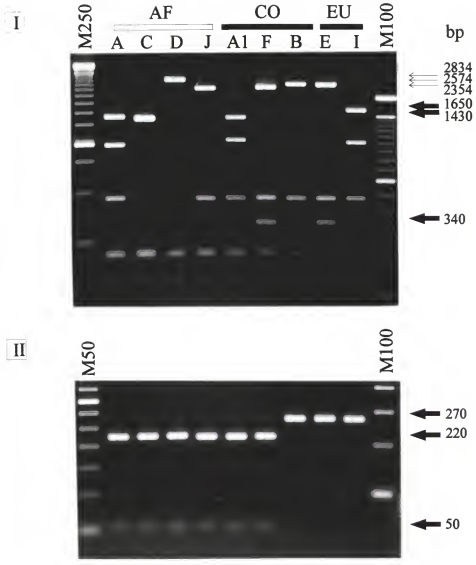


Figure 2-2. Autoradiograph of alleles revealed by *Alu*I digestions of probe P227 in African and European drone samples as revealed by Southern blot analysis. Numbers on top of each lane are the individual alleles. Fragment sizes are indicated in kilobase pairs. Alleles 1, 2, 3, 4, 5, and 6 are African-specific and alleles 7, 8 and 9 are European-specific (numbers were used in this case to distinguish alleles found in Southern blots to those found with PCR-RFLP analyses. These alleles do not necessarily correspond to the PCR-RFLP sub-alleles).

Figure 2-3. Restriction digest profiles of P227-S3xt (I), and P227-S3ter (II) with *Alu*I. AF=African sub-alleles, CO=common sub-alleles and EU=European sub-alleles. M50, M100, and M250 correspond to molecular size standard, 50, 100, and 250bp ladders, respectively (Lifetechnologies, GIBCO BRL, Gaithersburg, MD). Fragment sizes, in base pairs, are shown to the right. DNA samples amplified were all from drones. African-predominant sub-alleles are characterized by fragments of 1430bp (*Alu*I-*A*, and *C*) and 2834bp (*Alu*I-*D*). *Alu*I-*J*, an African-specific sub-allele, does not have a unique fragment and is distinguished from the others by its unique banding pattern. The European-predominant sub-allele *Alu*I-*I* is distinguished from the others by a 1650bp fragment and *Alu*I-*E* by its unique banding pattern. *Alu*I-*E* is an *A.m. caucasica*-specific sub-allele. Sub-alleles *Alu*I-*B* and *Alu*I-*J* are difficult to distinguish in RFLP patterns of P227-S3xt (I) but are clearly distinguishable in *Alu*I digests of P227-S3ter (II).


I
M250
AF
CO
EU
A C D J A1 F B E I
M100
bp
2834
2574
2354
1650
1430
340
II
M50
M100
270
220
50

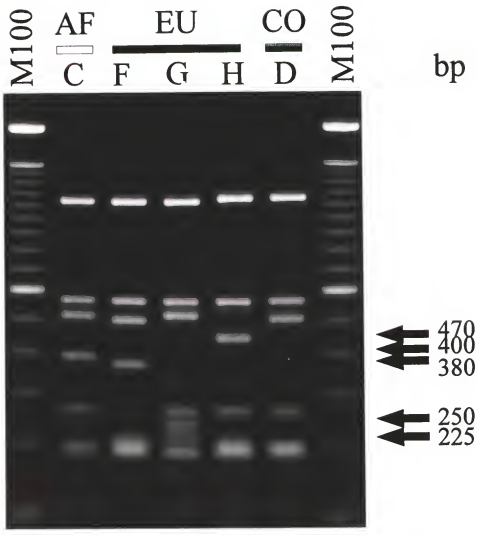


Figure 2-4. Restriction digestion profile for P227-S3xt with *Hinf*I. AF, CO, EU and M100 are as in Figure 2-3. European-specific sub-alleles are characterized by fragments of 380 (*Hinf*I-*F*), 250 and 225 (*Hinf*I-*G*), and 470 (*Hinf*I-*H*)bp. Allele *Hinf*I-*F* is east European-specific, *Hinf*I-*G* is west European-specific. *Hinf*I-*H* is found only in the U.S.A samples.

CHAPTER 3
DETECTION OF NUCLEAR RFLPs IN AFRICAN AND EUROPEAN HONEY BEE DNA AMPLIFIED WITH STANDARD AND LONG PCR.

Introduction

In Chapter 2, a locus was described with DNA restriction fragment length polymorphisms (RFLPs) that distinguish African and European honeybees. The RFLPs were sought and found using cloned probes in Southern blot analyses. Then, to facilitate DNA tests, the region with the RFLP alleles was made analyzable using the polymerase chain reaction (PCR). Other loci with RFLPs distinguishing African and European honeybees and found in Southern blots have also been made PCR analyzable (Hall, 1998). This conversion is time consuming: the position(s) of the informative polymorphic site(s) in the clone must be ascertained, so that the polymorphic region can be subcloned and subsequently amplified. In another approach, regions of DNA are first amplified, using primer sequences obtained from anonymous clones, and the regions are then screened for RFLPs using a large collection of enzymes (Karl *et al.*, 1992; Karl and Avise, 1993). The short length of DNA amplified by standard PCR limits the amount of variation observed, and, thus, a large number of clones must be screened to find useful polymorphisms. With modifications of the PCR, regions of DNA of up to 35 kilobase pairs (Kbp) can now be amplified (Barnes, 1994; Cheng *et al.*, 1994a; Cheng *et al.*, 1994b). In this chapter, we describe markers that distinguish African and European bees found by both standard and long PCR approaches, thereby demonstrating their relative

effectiveness. As expected, screening for RFLPs in long amplified regions was more productive than in short regions. Associations between alleles generated by pair of enzymes at a locus were also determined which reinforce the African and European specificity of the markers found.

Materials and Methods

Honey Bee DNA Samples

Samples were drone larvae and pupae from South Africa and the United States (European) as described previously in Chapter 2. Drone samples were primarily used because of their simple haploid pattern of restriction fragments. Other samples came from Italy, ten workers (diploid females) from five different colonies of the east European subpecies *A. m. ligustica*, and from France, ten workers from five colonies of the west European subspecies *A. m. mellifera* (provided by B. Vaissiere, and J-M. Cornuet, INRA, France). The samples of European workers were larvae or pupae, whereas in the work reported in Chapter 2, the worker samples were adults. Better quality and quantities of DNA can be extracted from larvae or pupae (suitable for long PCR analysis). The worker larvae were used to determine the origin of European-specific alleles but not to determine allele frequencies.

DNA Isolation

DNA isolation was as described (Hall, 1990) with modifications to increase quality and decrease shearing of the DNA, enabling amplification of long fragments. Two phenol-chloroform extractions were used instead of one, and, in each extraction, the

samples were gently mixed for 30min. Wide bore pipette tips were used to avoid mechanical damage of the DNA. DNA isolations were also performed using anionic columns (QIAGEN, Valencia, CA) according to the manufacturer's recommendations. DNA concentration was determined by UV absorbance at 260nm, and samples were diluted in TE (10mM Tris-HCl (pH 8.0), 1mM EDTA) to a final concentration of 100ng/ µl.

Clone Selection.

Random pieces of honeybee DNA digested with *Xba*I and *Pst*I were cloned in the multicloning site of pGEM3Z (Promega, Madison, WI). Plasmids were transformed into DH5α competent cells (Lifetechnologies, Gibco BRL, Grand Island, NY) and screened on bacterial plates with X-gal and ampicillin for white colonies with inserts. Plasmid minipreps were obtained by the alkaline lysis method as described in Sambrook *et al.*, 1989. Insert sizes were verified by digesting the DNA minipreps with *Xba*I and *Pst*I and separating the fragments by agarose gel electrophoresis. Two libraries were established: clones with inserts in the range of 1-2 Kbp, for amplification with standard PCR; and clones with inserts of more than 5 Kbp, for amplification with long PCR. Clones were designated as XPS or XPL (*Xba*I and *Pst*I, short or long inserts) followed by a number. Five clones with short inserts and five clones with long inserts were randomly selected. About 400 bps from both ends of each insert were sequenced (Interdisciplinary Center for Biotechnology Research, ICBR, University of Florida). From these terminal sequences, primer sequences were selected using the computer software package OLIGO 5.0 (National Bioscience, Plymouth, MN). Primers were synthesized by Lifetechnologies and by the ICBR, University of Florida. The primer sequences for each amplified region are presented in Table 3-1.

PCR Amplification and Restriction Enzyme Digests

Amplification conditions were optimized for the primer pairs selected from each clone (Table 3-1). If no amplification was achieved using the first pair of selected primers, another pair was designed. Amplifications were performed in 50mM Tris-HCl pH 9.2, 2.0 mM $MgCl_2$, (1.5mM $MgCl_2$ was used for regions XPS-3 and XPL-2), 16mM $(NH_4)_2SO_4$, with 2.0 units of *Taq* DNA polymerase and 0.25 units of *Pwo*I DNA polymerase. Both polymerases were taken from a single stock mixture (Boehringer Mannheim, Indianapolis, IN). Region XPS-3 was amplified with one unit of *Taq* DNA polymerase without *Pwo*I DNA polymerase. Amplification reactions, for both standard and long PCR, were in 50µl total volume, containing 200µM of each dNTP and 250nM of each primer. Totals of 100ng and 500ng of template DNA were used for the short and long regions, respectively. To increase the specificity of the amplification and the yield of long fragments, wax mediated Hot Start PCR was used (Chou *et al.,* 1992). Briefly, for each long PCR reaction, 1/3 of the total amount of water, 2/3 of the total volume of buffer, the nucleotides and the primers were aliquoted into one 0.5ml thin walled microcentrifuge tube and overlaid with 35µl of liquid wax (MJ Research, Watertown MA). The mixture was heated at 75^{o}C for 5min and cooled to 5^{o}C for another 5min. Then 2/3 of the total amount of water, 1/3 of the total amount of buffer, the template DNA and the enzyme mix were added on top of the solidified wax layer, and the optimized PCR program for the selected pair of primers was run. This procedure substantially reduced the formation of primer dimers and increased the yield of PCR products. Hot Start PCR was needed only for the long amplifications. The short region and two long regions (XPL-2 and XPL-5) were amplified with standard PCR profiles (Table 3-1). An autosegmented profile (progressive increase of the extension time per cycle) was used for one long region XPL-1 (Table 3-1).

Screening for RFLPs

To screen for polymorphic sites, 5µl of PCR products from one African and one European sample were digested by the addition of 5µl of a restriction digest mixture containing 2 units of the respective restriction enzyme and restriction digest buffer to a final concentration of 1X, as suggested by manufacturers (for a final volume of 10µl). The restriction enzymes used were *Alu*I, *Ava*I, *Ava*II, *Bam*HI, *Bcl*I, *Bgl*II, *Bst*XI, *Dde*I, *Eco*RI, *Eco*RV, *Hae*III, *Hha*I, *Hinc*II, *Hin*fI, *Hpa*I, *Kpn*I, *Msp*I, *Nci*I, *Nde*I, *Nhe*I, *Nsi*I, *Nsp*I, *Pvu*II, *Sal*I, *Sau*3A, *Sau*96, *Sca*I, *Sma*I, *Spe*I, *Sph*I, *Sst*I, *Xho*I, and *Xma*I. Restriction fragments were separated by electrophoresis in 2% agarose gels, stained with ethidium bromide and visualized over UV light. After this first screening, the locus enzyme combinations (LECs) that generated different banding patterns between the African and the European sample were tested on a larger sample size consisting of 48 African and 70 European drones. Alleles were named with the clone followed by the enzyme and a letter designating the allele (e.g. XPL-1/*Alu*I-*A*). Composite alleles were determined for loci XPL-2 and XPL-5, whereby each individual drone was assigned a combination of letters corresponding to the individual alleles produced by each enzyme at a particular locus.

Estimation of Allele Frequencies and Associations.

Drones are haploid parthenogenetic progeny of diploid queens. Therefore, frequencies of both individual and composite alleles were estimated in African and European populations on the basis of queen genotypes obtained from analyzing their drone progeny. Two to six drones per colony were analyzed. If two different alleles were found in drones from the same colony (heterozygous queen), each allele was scored once. If only one allele was found, it was scored once, or twice if found in at least six

drones (98% probability the queen was homozygous) (Hall, 1990). Contingency tables with intralocus association frequencies of individual alleles produced by pair of enzymes was constructed, for example, alleles produced by *Ava*I and *Hae*III at locus XPL-2. Association frequencies in all contingency tables were tested for independence using Fisher's exact test with the function STRUC in GENEPOP v1.2 (Raymond and Rousset, 1995).

Results

Comparisons of RFLPs Obtained in DNA Amplified by Standard PCR and Long PCR.

Of the five clones selected for standard short PCR, three were successfully amplified, one could not be amplified, and one generated PCR products of the expected size but failed to amplify in 30% of the samples. Of the five regions selected for long PCR, three generated PCR products of the expected size and two failed to amplify, even when new sets of primers were designed. PCR product sizes ranged from 1.6 to 2.0 Kbp for the short regions and from 5.5 to 9.0 Kbp for the long regions. The total number of restriction enzymes tested ranged from 26 to 31 for both short and long regions. Between the initial pair of African and European samples, 5 of 78 LECs (6.4%) tested from the short regions generated different band patterns compared to 25 of 85 LEC's (29.4%) from the long regions. After a large number of samples were screened for those enzymes that revealed polymorphisms during the initial screening, only 1 of 78 of the initial number of LECs (1.3%) tested from the short regions revealed specific nuclear DNA RFLPs between African and European populations, compared to 7 of 85 LEC's (8.2%) from the long regions.

Informative RFLPs were found in one short region, XPS-3, with *Alu*I. More were found within the long regions: XPL-1 with *Alu*I, XPL-2 with *Ava*I, and *Hae*III, and XPL-5 with *Hae*III, *Dde*I and *Spe*I. Fifty individual alleles were generated, of which 22 (44%) were African-specific, 13 (26%) were European-specific, two (4%) were predominant in European samples, and 13 (26%) were common to both African and European populations. Of the 50 individual alleles, only seven came from the short region XPS-3. For loci XPL-2 and XPL-5, composite alleles were determined from the combination of individual alleles. Forty-three composite alleles were found of which 22 (51%) were African-specific and 19 (46%) were European-specific (Table 3-2).

Description of DNA Polymorphisms

Collective frequencies of African and European-specific individual and composite alleles for each locus in the United States and South Africa are presented in Table 3-3. XPS-3 is the only clone less than two Kbp in length analyzed in this study that generated distinguishing RFLPs between African and European samples. Detailed characterization of this polymorphic locus will be reported in Chapter 4. At locus XPL-1, *Alu*I generated eight alleles of which one was European-specific (*H*), one was European-predominant (*F*), and five were African-specific (*B*, *C*, *D*, *E* and *G*) (Table 3-2 and Figure 3-1). At locus XPL-2, *Ava*I, and *Hae*III revealed polymorphisms. *Ava*I generated six alleles of which one was European-specific (*F*), one was European-predominant (*D*), and two were African-specific (*A* and *C*) (Table 3-2 and Figure 3-2). *Hae*III generated twelve alleles of which six were European-specific (*G*, *H*, *I*, *J*, *K* and *L*) and three were African-specific (*D*, *E* and *F*) (Figure 3-2). At locus XPL-2, a total of 25 composite alleles were found of which 11 were African-specific and 14 were European-specific (Table 3-2). At locus

XPL-5, *Hae*III, *Dde*I and *Spe*I revealed polymorphisms. *Hae*III produced 10 alleles of which two were European-specific (*G* and *H*) and six were African-specific (*A*, *B*, *C*, *E*, *F* and *I*) (Table 3-2 and Figure 3-3). *Dde*I generated five alleles of which two (*B* and *E*) were European-specific and two were African-specific (*C* and *D*) (Table 3-2 and Figure 3-3). *Spe*I produced only two alleles corresponding to the presence and absence of one restriction site in the entire region (Figure 3-3). *Dde*I allele *B* and *Spe*I allele *B* may be west European-specific. From each of 5 colonies of east (*A.m. ligustica*) and west (*A.m. mellifera*) European bees, two workers were selected and tested (drones from Europe were not available for this study). *Dde*I allele *B* and *Spe*I allele *B* were found in all the west European bees but not in the east European or African bees.

At locus XPL-5, a total of 18 composite alleles were found of which 11 were African-specific and seven were European-specific (Table 3-2).

For each LEC, the diploid genotypic banding pattern was constructed for every pairwise combination of alleles that might be found in worker bees, and the pairs of alleles that would be indistinguishable from others were determined. Only two LECs (XPL-2 *Hae*III and XPL-5 *Hae*III) produced indistinguishable genotypes, probably as a result of the complex banding pattern and the large number of low molecular weight fragments generated with *Hae*III. For LEC XPL-2 *Hae*III, the following pairs of genotypes could not be distinguished: *CC* and *CL*, *IK* and *IL*, *AJ* and *AK*, and , *JK* and *JL*. For LEC XPL-5 *Hae*III, genotypes *JJ* and *DJ* could not be distinguished.

Allele Associations.

To determine the association of individual alleles generated by pairs of different enzymes at a locus, Fisher's exact tests were used on contingency tables. The absence of

independence between alleles generated by pair of enzymes at a particular locus was statistically significant for all five comparisons ($P<0.01$). For example, at locus XPL-2, drones with allele *Ava*I-*D* are likely to have allele *Hae*III-*G* (Table 3-2). High frequency European alleles were generally associated with each other. For example, for locus XPL-2, strong associations were found between *Hae*III-*G*, and *Ava*I-*D* (Table 3-2). Other associations between European alleles were found, but most corresponded to low frequency alleles. African-specific alleles were, in general, associated with common alleles and not with other African-specific alleles (Table 3-2).

Discussion

As an alternative to Southern blot analysis with cloned probes and as an enhancement of standard PCR-RFLP, we have implemented the long PCR approach in our search for nuclear DNA markers specific to African and European bees. As would be expected, and as shown by our results, increasing the size of the amplified fragment increases the probability of detecting RFLPs. From a total of seven LECs that generated informative polymorphisms between African and European samples, six came from long regions compared to one from a short region. The sizes of the long PCR products in this study were about the same as cloned probes used in earlier studies, and alleles were obtained that showed similar distributions in African and European populations (Hall, 1992a; McMichael and Hall 1996).

The amplification of long DNA fragments is more demanding than standard PCR. Higher quality and larger amounts of DNA are required. Amplification of long DNA fragments is more sensitive to the reaction conditions and may require the use of wax mediated Hot Start PCR to increase the specificity of primer annealing and the yield of PCR product. The chances of false primer binding increase as the size of the fragment to be amplified increases, and, therefore, longer primers are required to increase the

specificity of annealing. Autosegmented programs are recommended for the amplification of long DNA fragments, which require, after an initial number of cycles, the progressive increase of the extension time (usually 20 seconds per cycle). A subroutine system can be used instead, in which the extension time is increased 100 seconds every five cycles. This modification provides comparable results to the autosegmented programs and is an acceptable substitute for thermocyclers that do not have the programming capability to perform autosegmented profiles.

For the regions amplified in this study, a number of different fragment patterns were generated by single restriction enzymes and are referred to as individual alleles. The combinations of these different alleles found in individual drones are called composite alleles. A composite allele is African or European-specific if one or more of its individual alleles is specific and its other individual alleles from different enzymes are common. Thus, compared to individual alleles, a greater proportion of composite alleles is considered specific, but, because of the larger number, frequency data from the same number of individuals are not as accurate. As expected, a lack of independence was found between pairs of individual alleles from different enzymes at the same locus. In some cases, pairs or triplets of African or European-specific individual alleles were strongly associated, which reinforced the specificity of the composite alleles.

As employed in this study, haploid drones are valuable for initially identifying different RFLP patterns and deciphering the allelic relationships of restriction fragments. However the use of drones for determining allele frequencies in populations has some drawbacks. Drones are usually available only during prosperous times and not from all colonies. Alleles in drones from the same colony represent those of only one individual, the mother queen. Thus, sample sizes are, in effect, limited to the number of colonies. Because honey bee queens mate with a number of drones from different colonies (Adams *et al.,* 1977; Page, 1986), worker progeny are more representative of the local population,

although the queen genotype would still be, and that of some paternal drones may be, over-represented.

It was shown that screening regions of DNA amplified by long PCR is an effective approach for finding useful RFLP markers. As discussed later (Chapter 5) more narrow sections which contain the informative restriction sites were isolated from the long regions, so that they could be detected using standard PCR. The resulting simpler patterns make the genotypes of workers easier to identify, and the less demanding protocol facilitates the analysis of large numbers of samples. More samples and populations tested in the future should result in more robust determinations of the specificity and frequencies of both the individual and composite alleles.

Table 3-1. PCR profiles and primer sequences used for the amplification of short and long regions

Locus	Primer sequences (5'-3')	Denaturation		Annealing		Extension		Final Extension		Cycles
(Size in kbp)		To	t	To	t	To	t	To	t	
Short region										
XPS-3	CGTGCCGACATTGAAACATTCCTCCA	94	1:00	65	1:00	72	2:00	72	15:00	45
(1.3)	TAGCCTCTGGTTCGTAACATCGAA									
Long regions										
XPL-1	GGCGTCGGTAGGTGATGTGAGGCTGGTG	94	0:45	65	0:45	68	6:00	68	15:00	10
(9.0)	TGCTTGATGGTGGTGGTCGGTGAGGA									>10 (*)
XPL-2	CGATGGCAGGACAGGATTGACGATG	94	0:45	61	0:45	68	6:30	68	15:00	32
(6.0)	TCATAGCACGACGACGACCACTACGG									
XPL-5	ATCCGGGATGTGAATTTCTTCTGA	94	0:45	61	0:45	68	5:30	68	15:00	32
(5.5)	CAAGGATGCGCCGAGTGTA									

To=temperature in °C; t=time in min:sec.

Protocols had an initial denaturation step of 94° for 1:00 min for the short region and 94° for 2:00 min for the long regions

(*): Region XPL-1 was amplified 10 cycles with the protocol described in the table, followed by 5 subroutines of 5 cycles each consisting of 94° for 0:45 min, 65° for 0:45 min and extension times set at 7:30; 9:00; 10:30 and 12:00 min for each subroutine.

Table 3-2. Number of individual and composite alleles in African and European populations. Percentage values relative to each population are indicated in parentheses.

		Population	
Locus (Enzymes)	**Allele**	**South Africa**	**United States**
XPS-3 (*Alu*I)	A^*	4 (18)	
	B^*	8 (35)	
	C^*	1 (4)	
	D^*	1 (4)	
	E^*	3 (12)	
	G	5 (22)	17 (68)
	H	1 (4)	8 (32)
n		23	25
XPL-1 (*Alu*I)	B^*	4 (16)	
	C^*	3 (13)	
	D^*	3 (13)	
	E^*	1 (4)	
	G^*	1 (4)	
	$H^\dagger$		2 (7)
	$F^\ddagger$	1 (4)	21 (72)
	A	11 (46)	6 (21)
n		24	29
XPL-2 (*Ava*I-*Hae*III)			
	A^*-*A*	3 (12)	
	A^*-*B*	1 (4)	
	A^*-*C*	3 (12)	
	A^*-E^*	1 (4)	
	B-*C*	3 (12)	

Table 3-2 continued.

B-D[*]	1 (4)	
B-E[*]	1 (4)	
B-F[*]	2 (8)	
C[*]*-C*	2 (8)	
D[‡]*-C*	1 (4)	
E-C	1 (4)	
B-H[†]		1 (3)
B-I[†]		2 (6)
B-L[†]		2 (6)
D[‡]*-A*		1 (3)
D[‡]*-C*		1 (3)
D[‡]*-G*[†]		15 (45)
D[‡]*-H*[††]		3 (9)
D[‡]*-I*[†]		1 (3)
D[‡]*-K*[†]		1 (3)
E-I[†]		1 (3)
E-J[†]		1 (3)
F[†]*-G*[†]		1 (3)
B-A	5 (20)	1 (3)
B-B	1 (4)	3 (9)
n	25	34

Locus XPL-5 (*Hae*III-*Dde*I-*Spe*I)

A[*]*-A-A*	5 (25)
A[*]*-D*[*]*-A*	2 (10)
B[*]*-A-A*	1 (5)
C[*]*-A-A*	1 (5)
D-A-A	2 (10)

Table 3-2 continued.

D-C**-A*	1 (5)	
D-D**-A*	1 (5)	
E**-A-A*	4 (20)	
F**-A-A*	1 (5)	
I**-A-A*	1 (5)	
J-A-A	1 (5)	
D-B†*-B*†		1 (4)
G†*-A-A*		7 (25)
G†*-B*†*-B*†		2 (7)
G†*-E*†*-A*		2 (7)
H†*-A-A*		8 (28)
H†*-E*†*-A*		1 (4)
J-B†*-B*†		7 (25)
n	20	28

(*): Individual African-specific allele.
(†): Individual European-specific allele.
(‡): Individual European-predominant allele

Table 3-3. Summary of markers found in short and long regions.

Locus	Enzyme	Population	Number of colonies tested (N)	Possible number of alleles (2N)[a].	Number of alleles counted	Total number of different alleles counted	Number of specific alleles (collective frequencies in %)
XPS-3	*Alu*I	South Africa	19	38	23	7	5 (73)
		United States	21	42	26		0 (0)
XPL-1	*Alu*I	South Africa	17	34	24	8	5 (50)
		United States	21	42	29		2 (79) [b]
XPL-2	*Ava*I	South Africa	17	34	25	6	2 (40)
		United States	22	44	34		2 (68) [b]
	*Hae*III	South Africa	17	34	25	12	3 (17)
		United States	22	44	34		6 (83)
Composite		South Africa	17	25	34	25	11 (76)
		United States	22	34	44		12 (90)
XPL-5	*Hae*III	South Africa	15	30	20	10	6 (75)
		United States	18	36	28		2 (71)
	*Dde*I	South Africa	15	30	20	5	2 (20)
		United States	18	36	28		2 (46) [c]
	*Spe*I	South Africa	15	30	20	2	0 (0)
		United States	18	36	28		1 (36) [c]
Composite		South Africa	15	20	30	18	11(100)
		United States	18	36	28		7 (100)

(a): Correspond to 2 X the number of colonies, i.e, the queens' genotypes.
(b): Collective frequencies include European-specific and European-predominant markers.
(c): Markers with intermediate frequencies in the United States population and with high frequencies in west European honey bee subspecies *A.m.mellifera.*

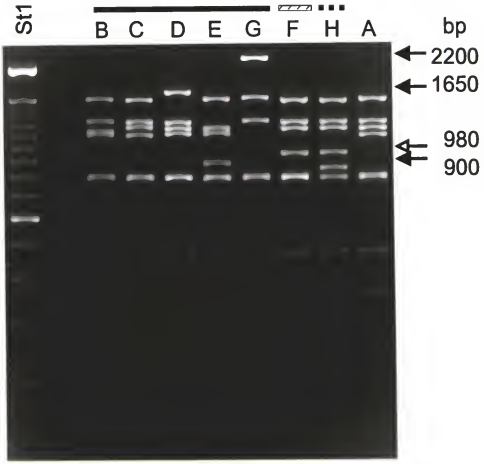


Figure 3-1. Alleles found in region XPL-1 with *AluI*. African-specific, European-specific and European-predominant alleles are indicated by a solid line, a dashed line, and a dotted line respectively. Bands characteristic of African-specific alleles are indicated by closed arrows. Bands characteristics of European-specific and European-predominant alleles are indicated by open arrows. European-specific allele *H* and European-predominant allele *F* both have a 980 bp fragment. African alleles are characterized by fragments of 1120bp (*B*), 1200bp (*C*), 1650bp (*D*), 900bp (*E*) and 2200bp (*G*). Alleles without an African or European-specific band are distinguished by their unique banding pattern. St1=Molecular size standard (100 bp DNA ladder. Lifetechnologies Gibco BRL).

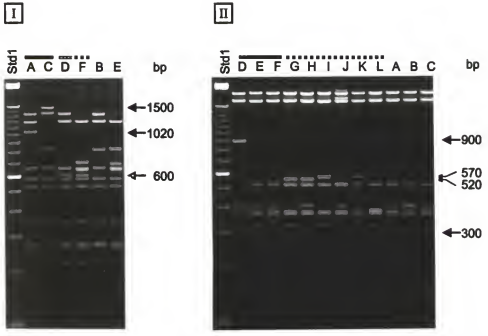


Figure 3-2. Alleles found in region XPL-2 with *Ava*I (I) and *Hae*III (II). Symbols are as indicated in Figure 3-1. (I) European-specific and predominant alleles found with *Ava*I are characterized by a 600bp fragment. African alleles are characterized by fragments of 1020bp (*A*) and 1500bp (*C*). (II) The European-specific alleles found with *Hae*III are characterized by the presence of a 550bp (*G* and *H*) or a 570bp (*I*, *J* and *K*) fragment. African-specific alleles are characterized by fragments of 900bp (*D*) and 300bp (*E*). African-specific allele *F* does not have specific fragments and is distinguished from the others by its unique banding pattern.

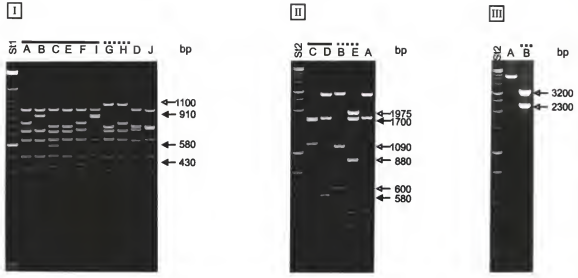


Figure 3-3. Alleles found in region XPL-5 with *Hae*III (I), *Dde*I (II) and *Spe*I (III). Symbols are as indicated in Figure 3-1. (I) European-specific alleles found with *Hae*III are characterized by a 1100bp fragment. African-specific alleles are distinguished by fragments of 910bp (*B*), 930bp (*I*), 580bp (*C*) and 430bp (*F*). (II) European-specific alleles found with *Dde*I are distinguished by fragments of 600 and 1090 bp for allele *B* and 1975 and 880 bp for allele *E*. African-specific alleles are characterized by a 1700 bp fragment for allele *C* and a 580 bp fragment for allele *D*. (III) *Spe*I produces two alleles of which one (*B*) is probably west European-specific. St2=Molecular size standard (1Kbp DNA ladder. Promega).

CHAPTER 4
CHARACTERIZATION OF LOCUS XPS-3

Introduction

This chapter describes the characterization of locus XPS-3, the only polymorphic locus of less than 2kbp found when comparing long PCR and standard PCR procedures to search for nuclear DNA markers (Chapter 3). RFLP analysis of locus XPS-3 digested with *Alu*I revealed alleles specific to or predominant in African and west European bees.

Materials and Methods

The samples used to determine allele frequencies for locus XPS-3 were as listed in Chapter 2. DNA was isolated from thoracic muscles of adult honey bees as described in Chapter 2. Primer sequences, amplification and digestion conditions, restriction cleavage mapping, frequencies and estimates of observed and expected heterozygosity, were described in Chapters 3 and Chapter 5 without further modification.

Results

*Alu*I digests of locus XPS-3 revealed seven alleles of which five were African-specific (*A*, *B*, *C*, *D*, and *E*), one was west European-predominant (*H*) and one was

common (Table 4-1 and Figure 4-1). African-specific alleles were collectively found at a 62.9% frequency in the South African samples. Allele *Alu*I-*H* was found fixed in the west European subspecies of *A.m.mellifera* and *A.m.iberica,* was found at low frequency in the east European subspecies *A.m.ligustica* and *A.m.carnica* (2.4% and 7.4%, respectively), and was found absent in the *A.m.caucasica* samples. *Alu*I-*H* was also found at low frequencies in the South African samples (2.3%).

A restriction cleavage map of locus XPS-3 is shown in Figure 4-2. All seven alleles resulted from the unique combinations of four polymorphic *Alu*I sites. African-specific alleles were characterized by fragments of 1050 (*A*), 250 (*B*), 550 (*C*), 940 (*D*) and, 900bp (*E*). The west European-predominant allele (*H*) was characterized by the presence of both a 100bp and a 150bp fragment (Figure 4-1). As with other polymorphic loci with African-specific alleles, genetic diversity values were higher in the South African population compared to the European populations (Table 4-1).

Discussion

Markers specific to African and west European groups of subspecies have been more difficult to find than to the east European groups. Generally, the number of African-specific alleles has been greater than that of east European-specific or predominant alleles resulting in a high genetic diversity in the African populations. However, the collective frequencies of the African alleles tend to be lower, reported here (Chapters 2, 3, 4 and 6) to average 40.5% compared to 80.0% for east European alleles. The 62.9% collective frequency for the African *Alu*I alleles at locus XPS-3 is among the

highest found in this study and is comparable to those for other polymorphic loci found previously (Hall, 1998: McMichael and Hall, 1996). The ability to discriminate African and west European bees from east European bees, the simple banding patterns, and the ease of standard amplification conditions make this polymorphic locus particularly useful.

Table 4-1: Allele frequencies, observed and expected heterozygosity (H_{obs} and H_{exp} respectively) and, collective African (C_{AF}), and west European (C_{WE}) allele frequencies in New World honey bee populations for *Alu*I digests of locus XPS-3.

	Populations						
Locus/Enzyme Allele	African		West European		East European		Other
	South Africa (*A.m.scutellata*)	United States	France (*A.m.mellifera*)	Spain (*A.m.iberica*)	Italy (*A.m.ligustica*)	Austria (*A.m.carnica*)	Russia (*A.m.caucasica*)
XPS3/*Alu*I							
(N)	132	67	13	15	42	26	19
A¥	0.1553						
B¥	0.3409						
C¥	0.0114						
D¥	0.1212						
E(*)							
G	0.3485	0.6716			0.9762	0.9231	1.0000
H§§	0.0227	0.3284	1.0000	1.0000	0.0238	0.0769	
C_{AF}	0.6288						
C_{WE}	0.0227	0.3284	1.0000	1.0000	0.0238	0.0769	
H_{obs}	0.6061	0.1493	0.0000	0.0000	0.0000	0.0769	0.0000
H_{exp}	0.7229	0.4411	0.0000	0.0000	0.0465	0.1420	0.0000

(*): Allele found in one drone but not found in workers.

$C_{(AF)}$: Collective frequencies of African-specific alleles.

$C_{(WE)}$: Collective frequencies of west European-predominant alleles

¥ : African-specific allele.

§§ : West European-predominant allele.

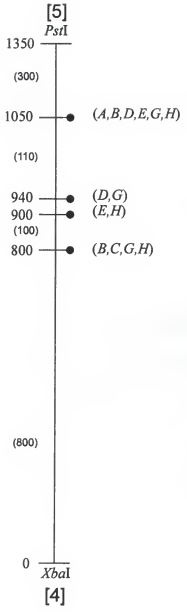


Figure 4-1: Restriction cleavage map of the polymorphic *Alu*I sites at locus XPS-3. Positions of the sites are indicated to the left (in bp). Letters in parentheses to the right, are the alleles that have the restriction sites. Numbers in brackets correspond to the primer sequences used to amplify locus XPS-3 (See Chapter 3). Numbers in parentheses indicate the sizes, in bp, of the fragments generated.

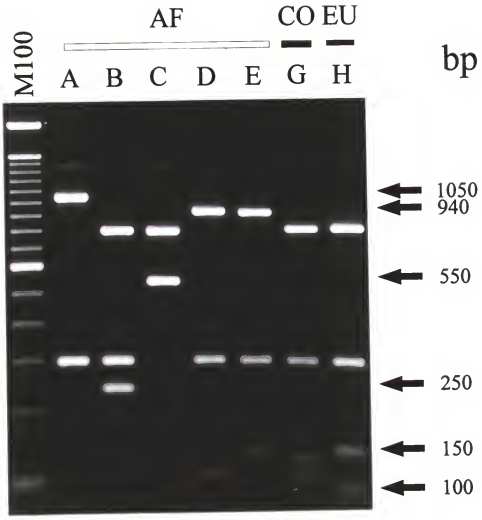


Figure 4-2: RFLP patterns of locus XPS-3 digested with *Alu*I. M100: Molecular size standard. African (AF) and European (EU) specific alleles are marked with an open and closed line respectively. Arrows to the right indicate fragment sizes, in bp, characteristic of each allele.

CHAPTER 5
CONVERSION OF LONG PCR MARKERS TO STANDARD PCR FORMAT

Introduction

Nuclear DNA-RFLPs that distinguish groups of African and European honey bee subspecies have been successfully found by amplifying and digesting anonymous regions of DNA with a number of restriction endonucleases (Chapter 3). Seeking informative RFLPs in long regions, amplified by modified procedures, has been more productive than in the short regions amplified by standard PCR. However, several problems remain concerning the amplification of large DNA fragments and the analysis of RFLP patterns of the amplifiable PCR products. First, high quality DNA is required. DNA from poorly preserved samples is difficult and usually impossible to amplify. Second, as the length of the amplifiable DNA region increases, usually the number of alleles increases, and their restriction fragment band patterns become more complex. The genotypes of diploid individuals, that is workers and queens in bees, become more difficult to determine. Therefore, conversion of long PCR RFLP markers to a standard PCR format (amplification of fragments of up to 2kbp) facilitates the amplification of DNA samples and the interpretation of band patterns. Conversion of RFLP markers found with probes on long PCR to a standard PCR format requires subcloning into sections and

amplification of the sections containing the informative polymorphic sites. This process is easier for polymorphisms found in amplified regions compared to those found with probes.

In this chapter, I discuss the conversion of the long PCR markers, previously described in Chapter 3, to a standard PCR format, and the characterization of these markers using a larger sample size of worker honey bees from different populations.

Materials and Methods

Honey Bee DNA Samples, Restriction Cleavage Maps, Subcloning, and DNA Amplification.

Samples of adult honey bee workers came from South Africa, Europe, and North America as previously described in Chapter 2. Conversion of long PCR to a standard PCR format was done for three loci, each of which have RFLP alleles generated with one to three enzymes: XPL-1 with *Alu*I, XPL-2 with *Hae*III and *Ava*I, and XPL-5 with *Hae*III, *Dde*I and *Spe*I. Restriction cleavage maps were obtained for the informative polymorphic sites and of other flanking enzyme sites, which enabled the regions containing the polymorphic sites to be subcloned and subsequently amplified using standard PCR procedures. Maps were generated using the procedure reported by Her and Weinshilboum (1995). This approach is illustrated in Figure 5-1 for two hypothetical alleles. For this mapping procedure, 40µl of PCR products from samples corresponding to each allele at a locus were partially digested with two units of the respective enzyme in 1X restriction digest buffer and water to a final volume of 50µl. Partial digestions were

obtained by aliquoting 13µl of the digestion mixture in a tube containing 12µl of 0.5M EDTA and 5µl of electrophoresis loading buffer (20% Ficoll, 0.5% bromophenol blue Tris-EDTA pH 7.0) at 2.5, 7.5, 15.0 and 30.0 minutes after initial incubation at 37°C. Partially digested products were electrophoresed at 4V/cm for 14-16 hours in a 2% agarose gel and transferred to a nylon membrane (Genescreen plus) by capillary blotting (Sambrook *et al*., 1989). The oligonucleotide primers used to amplify each region served as probes. Primers were end-labeled in 25µl with 50ng of primer, 1 unit of T4 polynucleotide kinase, 1X exchange reaction buffer and 25µCi of γ^{32}P (Lifetechnologies, Gibco BRL, Gaithersburg, MD). Labeling reactions were performed at 37°C for 30min. Unincorporated nucleotides and other inpurities were removed using anionic columns (Nucleotide removal kit, Qiagen, Palo Alto, CA). Blots were dried, then prehybridized for two hours, and hybridized for three hours at 37°C with ExpressHyb hybridization solution (Clonetech, Palo Alto, CA). Following hybridization, the blots were given two washes of 200ml each of 2XSSC + 0.1%SDS and two washes of 200ml each of 0.1XSSC + 0.5%SDS at room temperature (Sambrook *et al*., 1989). Each wash was for 30 min. Membranes were sealed in plastic bags and placed against X-ray film, held between intensifying screens at –70°C overnight. For each locus, hybridization was done first with the forward primer. Blots were stripped and then hybridized with the reverse primer. By using both forward and reverse primers as probes, the position of the restriction sites relative to each end of the section can be determined. This confirms the position of the site and allow for more accurate mapping of the sites, especially at the termini of the sections. Sites for enzymes found in the polycloning section of the plasmid vector pGem3Z (Promega Corporation, Madison, WI) were also mapped within the

sections. For subcloning of honey bee DNA, each region was first digested with the enzymes having sites both within the polycloning section of pGem3Z and that flanked the polymorphic sites. Digested fragments were separated in a 1% agarose gel and purified (Agarose gel extraction kit, Qiagen, Palo Alto, CA). Purified fragments were then subcloned in the plasmid vector pGem3Z by conventional cloning procedures (Sambrook *et al.*, 1989). Sequences for the terminal 400 bases of each subclone were obtained and used to design primer sequences with the software package OLIGO 5.0 (National Bioscience, Plymouth, MN). Amplifications were optimized for each region for magnesium ion concentration and cycling parameters (annealing temperature and time). To check for the presence of polymorphic sites in each region, 5µl of PCR products were digested with the respective enzyme as explained in Chapter 2. Digested PCR products were electrophoresed at 3V/cm in a 2% agarose gel and visualized by ethidium bromide staining over UV light.

Allele Determination and Data Analysis

Alleles found for the regions amplified by long PCR were named using the enzyme name followed by a letter (See Chapter 3). Amplifiable sections were identified using numbers and sometimes an abbreviation based on their position relative to the entire polymorphic locus (for example, XPL-2S1 for section 1 of locus XPL-2, and XPL-2S1int for section 1 internal of locus XPL-2). RFLP analysis of subclones from the original polymorphic loci results in simpler allelic band patterns, but because usually only a subset of the original restriction sites are present, distinction between some alleles, originally found with long PCR, was not possible. If an allele could be distinguished

from the others, it retained the same description used for the long PCR alleles. Otherwise, groups of alleles with the same banding patterns in a subcloned section were named with an "*X*" followed by a number. For locus XPL-2 and XPL-5, one to three sections were amplified to determine the allele identity: two sections for LECs XPL2 *Ava*I and XPL-5 *Hae*III and three sections for LEC XPL-5 *Dde*I. When two or three sections were used, the original allele identity was determined by the combination of band patterns found for the different sections. For each locus and population, the frequency of each allele and genotype was determined and used to calculate expected and observed heterozygosity (Nei, 1978 and Nei, 1987).

Results

For each polymorphic locus first amplified with long PCR (XPL-1, XPL-2 and XPL-5), one to three different sections were reamplified. These sections contained the informative polymorphic sites of a particular restriction enzyme and were of sizes amplifiable by standard PCR protocols: 0.48 kbp to 3.19 kbp (Table 5-1). Amplification conditions for all the sections within each locus are described in Table 5-1.

Characterization of Polymorphic Sections

Locus XPL-1. As detailed in Chapter 3, RFLP analysis of locus XPL-1 revealed polymorphisms with the restriction endonuclease *Alu*I. Eight alleles were found initially of which one was European-specific and one was European-predominant. The original European alleles were all characterized by the presence of a 980 bp fragment. To

subclone the section with the polymorphic sites responsible for this European-predominant fragment, locus XPL-1was divided in three different sections (Figure 5-2). As shown in Figure 5-2, these sections are defined by the position of an *Eco*RI site at position 3146 and a *Bam*HI site at position 5852. Positions of the *Alu*I sites that generated the different alleles were also mapped (Figure 5-2). The *Alu*I site at position 4990 is present in the European-specific and predominant alleles and absent in the other alleles initially found with long PCR. Therefore, the section containing this polymorphic site (XPL-1S2, between the *Eco*RI and *Bam*HI sites) was subcloned and amplified for RFLP analysis using standard PCR protocols. To verify the presence of the *Alu*I site in the amplifiable XPL-1S2, DNA from drone samples containing the different alleles, initially found using long PCR, were amplified and digested with *Alu*I (Figure 5-3). From these digests, three sub-alleles were found, compared to eight alleles in the entire region; one African-specific (*G*), one common (*X1* which correspond to the original alleles *Alu*I-*A*, *B*, *C*, *D* and *E* found using long PCR) and one European-predominant (*X2* which correspond to the original alleles *Alu*I-*F* and *H*). Sub-allele *G* is characterized by a 2050bp fragment. The European-predominant sub-allele is characterized by fragments of 150 and 700 bp (Figure 5-3).

Frequency data of the three different sub-alleles found in XPL-1S2 is presented in Table 5-2. The African-specific sub-allele *G* was not found in any of the worker samples but was found at low frequency in the African drone samples (Chapter 3). Frequencies of the European sub-allele, *X2*, were high in the east European samples (97.4%, and 92.5% in *A.m.ligustica*, and *A.m.carnica,* respectively) and, in the near east samples (70.0% in *A.m. caucasica*). Sub-allele *X2* was also found in high frequencies in the U.S.A samples

(78.7%). In the west European samples, sub-allele *X2* was absent in *A.m. iberica* but present at a low frequency in *A.m.mellifera* samples (4.5%) and in the South African samples (3.0%). Therefore, sub-allele *X2* was predominantly east European.

Locus XPL-2. Analysis of locus XPL-2 revealed polymorphisms with *Ava*I and *Hae*III (Chapter 3). Two sections from this locus were amplified. The first section, XPL-2S1int, is an 830bp fragment within section 1 located between the terminal *Pst*I site and the *Bam*HI site at position 1280. A second section, XPL-2S2, is a 3195bp fragment located between the *Eco*RI site at position 2765bp and the *Sph*I site at position 6230 (Figure 5-4). Sub-alleles for *Ava*I were identified by amplifying both XPL-2S1int and XPL-2S2 (Figure 5-4 and Table 5-3). In section XPL-2S1int, three sub-alleles could be distinguished: *Ava*I-*A* (African-specific), characterized by an 830bp fragment, *Ava*I-*X2* (European-predominant) characterized by a 600bp fragment, and*Ava*I-*X1* (common) (Figure 5-5 I). XPL-2S2 has three different sub-alleles: *Ava*I-*C* (African-specific), characterized by a 1505bp fragment, and *Ava*I-*X1* and *X2* (common) (Figure 5-5 II). Identification of the six alleles originally found with long PCR was possible by looking at the unique sub-allele combinations found in both polymorphic section: XPL-2S1int and XPL-2S2. Combinations *A*:*X1*, *X1*:*C*, *X2*:*X1*, *X2*:*X2*, *X1*:*X1*, *X1*:*X2* (first letter is the sub-allele found in XPL-2S1int and the second letter is the sub-allele found in XPL-2S2) correspond to the alleles *Ava*I-*A*, *C*, *D*, *F*, *B* and *E*, respectively revealed using long PCR (Table 5-3). For *Hae*III, only one section was amplified (XPL-2S1int). Three alleles of the original 12 alleles found with long PCR were clearly distinguishable: *Hae*III-*E*, *F* (African-specific) and *Hae*III-*K* (European-specific). The remaining alleles were

grouped by four banding patterns of which two, *Hae*III-*X2* and *X3*, were European-specific (Figure 5-6) corresponding to four of the original six European-specific alleles seen with long PCR (*Hae*III-*G*, *H*, *I* and *J*). The European-specific sub-alleles *Hae*III-*X2*, *X3* and *K* were characterized by a 520bp fragment (Figure 5-6). Original alleles *Hae*III-*D* (African), *L* (European) and *C* (common) were seen as *X1*, and alleles *Hae*III-*A* and *B* as *X4*.

Sub-allele frequencies for alleles revealed by *Ava*I and *Hae*III in African and European populations are shown in Table 5-2. Alleles *Ava*I-*D* and *F* were found collectively at frequencies of 77.1% and 97.6% in the east European subspecies *A.m.ligustica* and *A.m.carnica* respectively, were absent in west European samples and present at very low frequencies in *A.m.scutellata* (South African samples) (1.1%) and in *A.m.caucasica* (near-East) (27.5%). These alleles were found at intermediate frequencies in samples from the United States (65.6%). African-specific alleles *Ava*I-*A* and *C* were found at a 51.1% collective frequency in the *A.m.scutellata* samples and were absent from the east and west European samples. However, *Ava*I-*A* was found at low frequencies in the U.S.A samples (3.7%). Sub-alleles *Hae*III-*X2*, *X3* and *K* were found at collective frequencies of 76.2, 94.2 and 92.5% in *A.m.ligustica, A.m.carnica* and *A.m.caucasica* samples, respectively. *Hae*III-*K* was found only in *A.m.caucasica* (30.0%) and, therefore, appeared to be specific to this subspecies. These sub-alleles are east European-specific, found to be absent from west European and the South African samples. African-specific sub-alleles *Hae*III-*E* and *F* were found at a low frequency in the South African samples (17.1%) and were absent in all the European and the U.S.A samples.

Locus XPL-5. Analysis of locus XPL-5 revealed polymorphisms with the enzymes *Hae*III, *Dde*I, and *Spe*I (Chapter 3). Locus XPL-5 was divided into four sections. The first section, XPL-5S1 is a 1499 bp fragment located between the *Hinc*II sites at positions 630 and 2129. The second section is a 900 bp fragment located between the *Hinc*II sites at positions 3029 and 2129. The third section is a 1501 bp fragment located between the *Hinc*II sites at position 4530 and 3029, and the fourth is a 900 bp fragment located between the *Xba*I site at position 5430 and the *Hinc*II site at position 4530. To detect the polymorphic *Hae*III site, two sections were amplified: section 1 extended (XPL-5S1xt), a 1980bp fragment located within section 1 between the *Pst*I site and the *Hinc*II site at position 2129, and section 1 terminal (XPL-5S1ter) a 485bp fragment that overlaps the *Hinc*II site near section 2 and the polymorphic *Hae*III site in XPL-5S1xt at position 1880 that is absent in the east European bees (Figure 5-7). Both sections were amplified which subdivided the complicated banding patterns originally found with *Hae*III. From the original ten alleles detected using long PCR, five sub-allele groups were revealed (*Y1*, *Y2*, *Y3*, *C* and *F*)(resulting from the combination of banding patterns found in both XPL-5S1xt and XPL-5S1ter) of which three are African-specific (*Y2*, *C* and *F*) and one is European-predominant (*Y3*) (Figure 5-8 I and II, and Table 5-2). The "Y" notation has been used in this case to avoid confusion with the conventional "X" notation used previously for groups of banding patterns. In this particular case, groups of banding patterns were found in both sections and were named with an "X" followed by a number. The final sub-allele designation is designated with a "Y". The African-specific sub-alleles are characterized by fragments of 974, 600 and 450bp respectively (in digests

of XPL-5S1xt) and the European-predominant sub-allele (*Y3*) is characterized by a 485bp fragment detected in digests of XPL-5S1ter (Figure 5-8 II).

Sub-allele *Y3* is East European-predominant, found at a frequency of 92.9 and 98.1% in *A.m.ligustica* and *A.m.carnica,* respectively, at 100.0% in *A.m.caucasica,* and at intermediate frequencies, 71.9%, in the USA samples. Sub-allele *Y3* was found at a 1.1% frequency in *A.m.scutellata* (South African).

Identification of *Dde*I alleles (*Dde*I *A*, *B*, *C*, *D* and *E*) was done by amplifying three sections of locus XPL-5: section 1 extended, section 2 (XPL-5S2), a 715 bp fragment located between the *Hinc*II sites at positions 2129 and 3029, and section 4 (XPL-5S4), a 935 bp fragment located between the *Hinc*II site at position 4530 and the terminal *Xba*I site (Figure 5-9). All the alleles found with long PCR (*Dde*I *A*, *B*, *C*, *D* and *E*) could be identified by the banding patterns detected in one or a combination of two or more of the three sections. Polymorphisms in section XPL-5S1xt distinguishes allele *B* from the others, section XPL-5S2 distinguishes alleles *C* and *E* and, section XPL-5S4 distinguishes allele *D*. Allele *A* is distinguished from the others by its banding pattern combination in the different sections (XPL-5S1xt, S2 and S4) (Figure 4-9). Allele *B* is west European-predominant found fixed in both *A.m.mellifera* and *A.m.iberica*, at 4.8 and 11.5% frequencies in the east European samples of *A.m.ligustica* and *A.m.carnica,* respectively, and, at a frequency of 5.0% in the *A.m.caucasica* samples. Allele *E* is east European-specific, found at a low frequency in *A.m.ligustica* (14.3%) and at an intermediate frequency in *A.m.carnica* (48.1%), and absent in the African and west European samples (Table 5-2).

Polymorphisms revealed by *Spe*I were detected by amplifying one section (XPL-5S3) of locus XPL-5. Two sub-alleles were detected resulting from the presence (sub-allele *B*) or absence (sub-allele *A*) of one restriction site at position 3290. Sub-allele *B* was European-specific, predominantly found in the west European subspecies: in *A.m.mellifera* at a 73.1% frequency and in *A.m.iberica* at a 33.3% frequency. Sub-allele *B* was also found at low frequencies in the east European subspecies *A.m.ligustica* (1.2%) and *A.m.carnica* (15.4%), and absent in *A.m.caucasica* (Table 5-2).

Discussion

Making long PCR markers analyzable by standard PCR has facilitated their use for population genetic studies. Amplification of small regions of DNA do not require high quantity and quality of DNA as with long PCR. Because the sizes of the amplified regions are reduced, the resulting band pattern of the alleles, initially identified using long PCR, are simplified. Thus, the genotypes of workers and queen (diploid) honey bees can be more easily determined. The quality of DNA suitable for long PCR can be consistently obtained when extracted from entire pupae but not from thoracic muscles of adult bees. The difficulty of obtaining such high quality of DNA from adults restricts the applicability of the long PCR markers. With the conversion of these markers to a standard PCR format, DNA from adult honey bee thoracic muscles can be easily analyzed. Moreover, the simplicity of the banding pattern and the small size of the fragments amplified make it easier to analyze diploid individuals. Thus, these

modifications make a large number of workers from different populations amenable for analysis.

To determine the genotypes of worker honey bees, one to three sections had to be amplified, and patterns obtained in each section were combined to obtain a final allele identification. In testing unknown samples, usually, identification of all the original alleles was possible. However, sometimes when a large number of alleles per locus was found in the population, the identification of both alleles in an individual worker using the standard PCR modification was not possible, because the phase (cis versus trans) of the polymorphisms relative to each other cannot be determined (as described in Chapter 2). For example, in the locus enzyme combination XPL-2/*Ava*I, the identification of each allele is done by amplifying two sections: XPL-2S1int and XPL-2S2. If a worker's genotype for XPL-2S1int is *X1/X2* and for XPL-2S2 is *X1/X2*, two possibilities exist: a cis configuration on homologous chromosomes will be *X1X1* on one chromosome (Allele *Ava*I-*B*) and *X2X2* on the homologous chromosome (Allele *Ava*I-*F*) (Table 5-3); a trans configuration will have *X1X2* (Allele *Ava*I-*E*) on one chromosome and *X2X1* (Allele *Ava*I-*D*) on the homologous chromosome. Thus, the individual could be either *B*/*F* or *E*/*D*. In this example, both alleles *B* and *E* are common alleles and, *D* and *F* are European-predominant, therefore, an incorrect identification of these alleles would not affect the collective frequencies of the European markers. These ambiguous cases were rarely found in our samples, and, typically, most of the alleles could be correctly identified.

As in any PCR-based method, erroneous readings can also be attributed to preferential amplification of certain alleles or to the presence or artifactual or

heteroduplex DNA (Anglani *et al*., 1990; Arnheim and Erlich, 1992; Hare *et al*., 1996). Partial restriction digestions can also confuse allele identification.

The methodological and analytical simplifications of the markers found with long PCR procedures greatly increase their value for both research and regulatory purposes. These genetic markers are thus an important addition to the collection of markers becoming available, which can be used to monitor the spread of the African bees in the United States.

Table 5-1: Amplification profiles and oligonucleotide primers used to amplify the polymophic sections of loci XPL-1,XPL-2 and XPL-5.

Locus/Region	Primer sequences (5'-3')	Size (Kbp)	Annealing		Extension		Cycles	Mg^{2+} ions (mM)
			(°C)	(min:sec)	(°C)	(min:sec)		
XPL-1S2	[12] TGACGTCTTCATCCACGAGT [13] ACCAACCTGAAGGCGAATTT	2.15	63	00:45	72	02:00	37	1.0
XPL-2S1int	[14] GGCGTCCAGGTAACCGTCTCC [15] CGGTTGGAGGCGAACGGAAA	0.83	65	00:45	72	01:00	35	1.0
XPL-2S2	[16] CCTCGACGTCGCTGGATGTTGC [17] CGCCACAAACGAGCCACCCCTACT	3.19	60	00:45	68	03:30	33	1.0
XPL-5S1xt	[18] GAATCCAAGGATGCGCCGAGTGTAT [19] GCGGGCTGTGAGAGGTGCATCTAAT	1.87	64	00:45	72	02:00	35	1.0
XPL-5S1ter	[20] GAAATCAATGGCCGGGGGACTCT [21] CATCCATGAAAGCACAATCGGAACG	0.48	65	00:45	72	00:45	35	1.5
XPL-5S2	[22] GTCGCGTTCCGATTGTGCTTTCAT [23] GCCCCATTTTTGCATGACGTATCG	0.71	65	00:45	72	00:45	35	1.0
XPL-5S3	[24] GATAAAGCGATTGGGAATGC [25] TGCTTCGTCGAAAGAGACACTT	1.28	55	00:45	72	2:30	40	1.5
XPL-5S4	[10] ATCCGGGATGTGAATTTCTTCTGA* [26] TCGAGCGTTAATTAGCTGACTGGT	0.93	60	00:45	72	01:00	37	1.5

Sections were amplified using an initial denaturation step of 94°C for 45sec. Sections were amplified in 50mM Tris-HCl pH 9.2, 16mM $(NH_4)_2SO_4$, with 200µM of each dNTPs, 250nM of each primer and 1.5units of *Taq* DNA polymerase in a final volume of 25µl.

*: Primer also used to amplify the locus XPL-5.

Table 5-2: Allele frequencies, observed and expected heterozygosity (H_{obs} and H_{exp} respectively) and, collective African (C_{AF}), east European (C_{EE}) and west European (C_{WE}) allele frequencies in New World honey bee populations.

Locus Enzyme	Populations						
	African	West European			East European		Other
	South Africa (*A.m.scutellata*)	United States	France (*A.m.mellifera*)	Spain (*A.m.iberica*)	Italy (*A.m.ligustica*)	Austria (*A.m.carnica*)	Russia (*A.m.caucasica*)
XPL-1/*Alu*I							
(N)	133	68	11	13	39	20	20
G							
X1	0.9699	0.2132	0.9545	1.0000	0.0256	0.0750	0.3000
X2[§]	0.0301	0.7868	0.0455		0.9744	0.9250	0.7000
C_{EE}	0.0301	0.7868	0.0455		0.9744	0.9250	0.7000
H_{obs}	0.0602	0.3088	0.0909		0.0513	0.1500	0.5000
H_{exp}	0.0583	0.3355	0.0868		0.0500	0.1387	0.4200
XPL-2/*Ava*I							
(N)	131	67	12	12	35	21	20
A[¥]	0.4122	0.0373					
C[*]	0.0992						
D[§]	0.0115	0.6567			0.7714	0.9286	0.2250
F[§]						0.0476	0.0500
B	0.4427	0.3060	1.0000	1.0000	0.2286	0.0238	0.6500
E	0.0344						0.0750
C_{AF}	0.5114	0.0373					

Table 5-2 continued.

C_{EE}	0.0115	0.6567			0.7714	0.9762	0.2750
H_{obs}	0.6183	0.5821			0.3429	0.1429	0.5500
H_{exp}	0.6229	0.4737			0.3527	0.1349	0.5188
XPL-2/*Hae*III							
(N)	120	66	13	15	42	26	20
E[*]	0.0208						
F[*]	0.1500						
X2[§]		0.5000			0.7024	0.8269	
X3[§]		0.1591			0.0595	0.1154	0.6250
K[§]		0.0076					0.3000
X1	0.3958	0.2045			0.2262	0.0385	
X4	0.4333	0.1288	1.0000	1.0000	0.0119	0.0192	0.0750
C_{AF}	0.1780						
C_{EE}		0.6667			0.7619	0.9423	0.9250
Hobs	0.4500	0.6667			0.5000	0.3077	0.4500
Hexp	0.6326	0.6662			0.4518	0.3010	0.5137
XPL-5/*Hae*III							
(N)	135	64	12	15	42	26	19
Y2[*]	0.0704						
C[*]	0.0296						
F							
Y4	0.0148						
Y3[§]	0.0111	0.7109			0.9286	0.9808	1.0000
Y1	0.8741	0.2891	1.0000	1.0000	0.0714	0.0192	
C_{AF}	0.1000						
C_{EE}	0.0111	0.7109			0.9286	0.9808	1.0000

Table 5-2 continued.

H_{obs}	0.1778	0.3281			0.1429	0.0385	
H_{exp}	0.2298	0.4110			0.1327	0.0377	
XPL-5/*Dde*I							
(N)	135	66	13	15	42	26	20
C							
D¥	0.2037						
B§§	0.0037	0.3712	1.0000	1.0000	0.0476	0.1154	0.0500
E§		0.1818			0.1429	0.4808	
A	0.7926	0.4470			0.8095	0.4038	0.9500
C_{AF}	0.2037						
C_{EE}		0.1818			0.1429	0.4808	
C_{WE}	0.0037	0.3712	1.0000	1.0000	0.0476	0.1154	0.0500
H_{obs}	0.3481	0.6212			0.3571	0.5385	0.1000
H_{exp}	0.3303	0.6294			0.3220	0.5925	0.0950
XPL-5/*Spe*I							
(N)	136	67	13	15	42	26	20
A	1.0000	0.6940	0.2692	0.6667	0.9881	0.8462	1.0000
B§§		0.3060	0.7308	0.3333	0.0119	0.1538	
C_{WE}		0.3060	0.7308	0.3333	0.0119	0.1538	
H_{obs}		0.4030	0.5385	0.6667	0.0238	0.3077	
H_{exp}		0.4247	0.3935	0.4444	0.0235	0.2604	

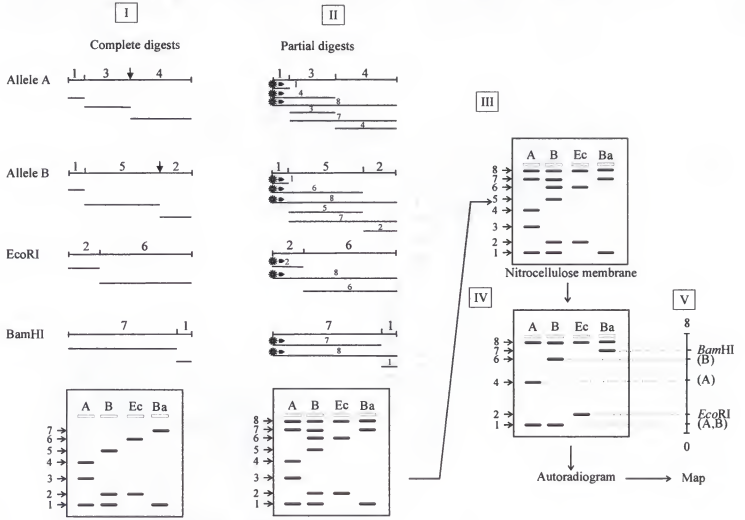
(§): East European-specific or predominant allele or sub-allele.
(§§): West European-predominant allele or sub-allele.
(¥): African-specific or predominant allele or sub-allele.
(*): Alleles found in very low frequencies and were not detected in worker samples but in one drone sample.
(N): Number of workers (2N= number of alleles).

Table 5-3. Allele determination as revealed by amplification of multiple sections using standard PCR procedures.

Locus Enzyme Combination (LEC)	Allele combinations found in each section	New Allele designation [1]	Long PCR alleles
XPL-2/*Ava*I	S1int/S2		
	A / X1	*A*¥	*A*¥
	X1 / C	*C*¥	*C*¥
	X2 / X1	*D*§	*D*§
	X2 / X2	*F*§	*F*§
	X1 / X1	*B*	*B*
	X1 / X2	*E*	*E*
XPL-5/*Dde*I	S1xt/S2/S4		
	X1 / C / X1	*C*¥	*C*¥
	X1 / X1 / D	*D*¥	*D*¥
	B / X1 / X1	*B*§§	*B*§§
	X1 / E / X1	*E*§	*E*§
	X1 / X1 / X1	*A*	*A*
XPL-5/*Hae*III	S1xt/S1ter		
	X1 / X1	*Y1*	*A*¥, *E*¥, *D*, *J*
	X2 / X1	*Y2*¥	*B*¥, *I*¥
	C / X1	*C*¥	*C*¥
	F / X1	*F*¥	*F*¥
	X1 / X2	*Y3*§	*G*§, *H*§

(1): Original allele found at locus XPL-5 that could not be distinguished by the combination of banding patterns from RFLP analysis of section 1 extended (S1xt) and section 1 terminal (S1ter) were grouped. Different group of alleles were named with a Y followed by a number.
(§): East European-specific or predominant allele.
(§§): West European-predominant allele.
(¥): African-specific or predominant allele.

Figure 5-1. Diagram describing the procedure used to produce the restriction cleavage maps and described by Her and Weinshilboum (1995). The procedure is illustrated using two hypothetical alleles (*A* and *B*) with two restriction sites for a hypothetical enzyme: one monomorphic and one polymorphic (the polymorphic site is indicated by an arrow). This locus also has one *Bam*HI and one *Eco*RI site that flank the polymorphic site for the hypothetical enzyme. The procedure starts by partially digesting PCR products from both alleles and separating these products by agarose gel electrophoresis. Following electrophoresis, the partially digested products are denatured and transferred to a nitrocellulose membrane. The nitrocellulose membrane is then prehybridized and hybridized with radioactively end-labeled primers (one primer) and exposed to X-ray film. Radiolabeled primers bind only to those fragments that contain the homologous primer sequences, that is that extend to the terminus of the sequence. The sizes of the fragments seen in the exposed autoradiogram therefore indicate the positions of the restriction sites relative to the position of the radiolabeled primer used. From this information, a map can be constructed. The positions of the polymorphic sites and the sections to subclone can be determined. I) Diagram showing the complete enzyme digestion for samples with alleles *A* and *B* and for both the *Eco*RI and *Bam*HI sites. Numbers indicate the fragment sizes. Expected electrophoresis patterns are show below. II) Partial digestions of PCR products from samples with alleles *A* and *B*, and the expected patterns shown after agarose gel electrophoresis (bottom). Numbers indicate fragment sizes. One of the primers used to amplify this locus is indicated with a point with an asterisk. Presence of this symbol indicates the fragments from the partial digestion to which this radiolabeled primer binds. III) Nitrocellulose membrane with transferred partial digested products. IV) Autoradiogram exposing bands, in which the radioactively labeled primer is bound. V) Resulting restriction cleavage map.



Agarose gel electrophoresis of complete and partial digests

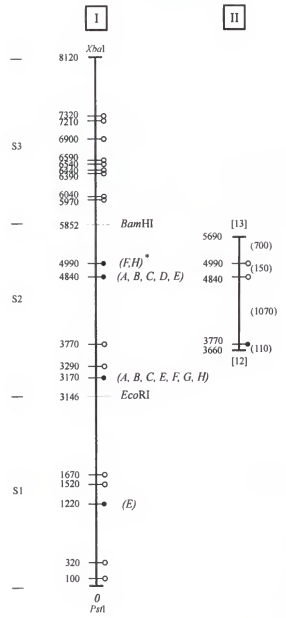


Figure 5-2. Restriction cleavage map of locus XPL-1 with the restriction enzyme *Alu*I. Numbers indicate the position of monomorphic (open symbols) and polymorphic (closed symbols) *Alu*I sites. Numbers in parentheses indicate sizes of fragments. Letters in parentheses are the alleles with the polymorphic sites. Numbers enclosed in brackets are of the primers used to amplify section 2 and their sequences are listed in Table 5-1. Locus XPL-1 was divided in three sections: section 1 (XPL-1S1) from the *Pst*I site at position 0 to the *Eco*RI site at position 3146. Section 2 (XPL-1S2) from the *Eco*RI site at position 3146 to the *Bam*HI site at position 5852, and section 3 (XPL-1S3) from the *Bam*HI site at position 5852 to the *Xba*I site at position 8120. I) Relative position of *Alu*I sites. II) Position of *Alu*I sites and the size of the resulting fragments in XPL1-S2. Indicated with an asterisk is the site that results in the east European predominant alleles *F* and *H*.

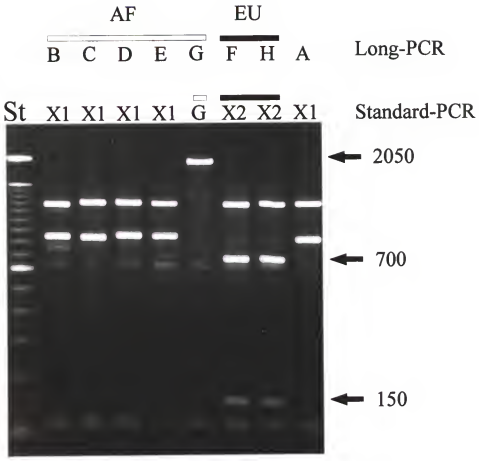


Figure 5-3. RFLP patterns revealed by *Alu*I digests of section 2 of locus XPL-1 (XPL-1S2). Three patterns are seen: the first pattern (*X1*) which includes most of the African alleles (*B*, *C*, *D* and *E*) and the common allele (*A*). The second pattern corresponds to the African-specific allele *G*, and, the third pattern (*X2*) corresponds to the European alleles (*F* and *H*). The pattern (*X2*) is characterized by the presence of a 700 and a 150bp fragment resulting from one restriction site located in the 850bp fragment found in pattern X1. St: Molecular size standard (100bp ladder, Lifetechnologies, Gibco BRL, Gaithersburg, MD). African (AF) and European (EU) alleles are marked with an open and closed line respectively. Letters in the first line are the allele designations initially found in long PCR amplifications and, just above the picture, the new sub-allele designations based on standard PCR amplification of this one region.

Figure 5-4. Restriction cleavage map of locus XPL-2 with the restriction enzymes *Ava*I and *Hae*III. Numbers and symbols to the left and right of each map correspond to the data for *Ava*I and *Hae*III, respectively. Circles and squares correspond to *Ava*I and *Hae*III sites respectively. (Polymorphic sites are indicated by a closed symbol and monomorphic sites are indicated with an open symbol). Letters enclosed in parentheses are the alleles with the polymorphic site. Numbers enclosed in brackets are of the primers used to amplify each region and their sequences are listed in Table 5-1. Numbers in parentheses indicate fragment sizes in base pairs. Locus XPL-2 is divided in three sections: Section 1 from the *Pst*I site to the *Eco*RI site (position 2765), section 2 from the *Eco*RI site to the *Sph*I site (position 6230) and section 3 from the *Sph*I site to the *Xba*I site (7130). The section between the *Pst*I and the *Bam*HI site is named section 1 internal (S1int) as it includes only half of the original section 1. I) Relative positions of restriction sites for *Ava*I (left), *Hae*III (right) and the enzymes used to subclone fragments of sizes amplifiable by standard PCR. II) Positions of both monomorphic and polymorphic *Ava*I and *Hae*III sites. Indicated with one asterisk are those polymorphic sites that generate European-predominant alleles, and with two asterisks those that generate African-specific alleles. For *Ava*I, the polymorphic sites located at positions 550 and 3835 generate African-specific alleles. The absence of the *Ava*I site at position 550 results in allele *Ava*I-*A* and the absence of the site at position 3835 results in allele *Ava*I-*C*. The presence of the site located at position 1150 results in the generation of the east European-predominant alleles *Ava*I-*D* and *F*. For *Hae*III, two sites are responsible for the African-specific allele: one at position 520 and the second at position 670. Alleles *Hae*III-*E* and *F* lack the former and *Hae*III-*F* lacks the latter. The absence of the *Hae*III site located at position 1120 results in the east European-predominant alleles *Hae*III-*G* and *H*. III) Regions amplified, using standard PCR protocols, that contain most of the polymorphic sites for *Ava*I and *Hae*III.

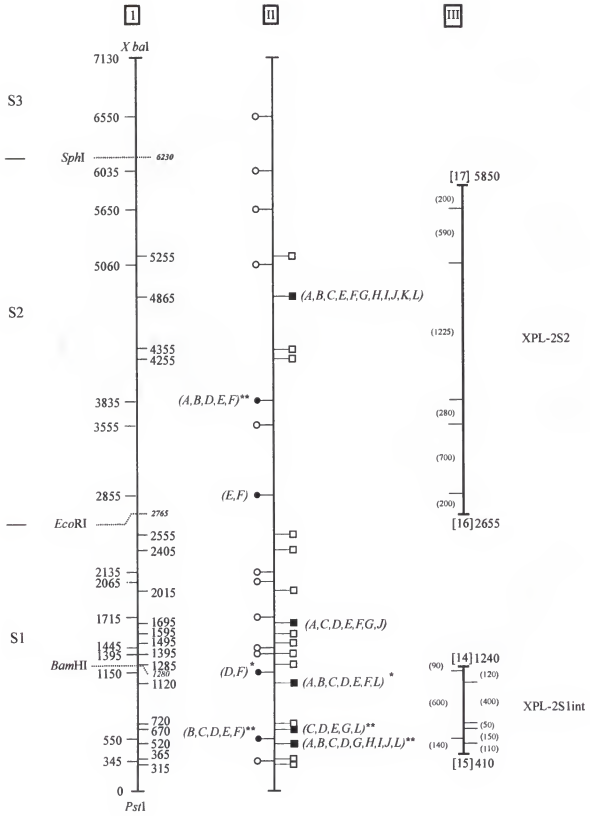
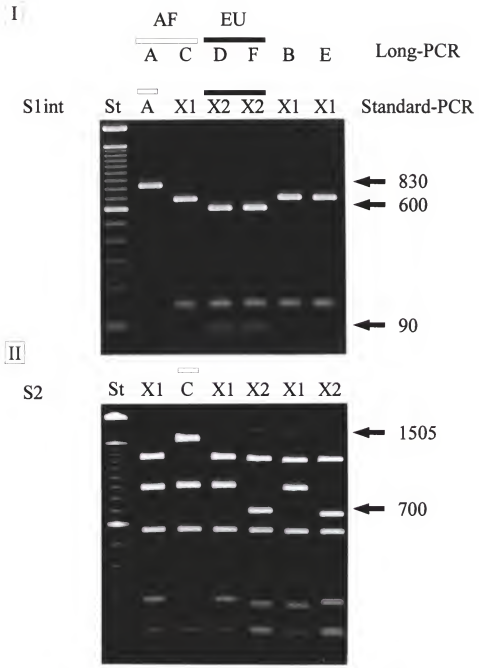

I
II
III
XbaI
7130
6550
SphI
6230
6035
5650
5060
5255
4865
4355
4255
3835
3555
2855
EcoRI
2765
2555
2405
2135
2065
2015
1715
1695
1595
1495
1445
1395
BamHI
1285
1260
1150
1120
720
670
550
520
365
345
315
0
PstI
S3
S2
S1
(A,B,C,E,F,G,H,I,J,K,L)
(A,B,D,E,F)**
(E,F)
(A,C,D,E,F,G,J)
(D,F)*
(A,B,C,D,E,F,L)*
(B,C,D,E,F)**
(C,D,E,G,L)**
(A,B,C,D,G,H,I,J,L)**
[17] 5850
(200)
(590)
(1225)
(280)
(700)
(200)
[16] 2655
XPL-2S2
[14] 1240
(90)
(120)
(600)
(400)
(50)
(150)
(140)
(110)
[15] 410
XPL-2S1int

Figure 5-5. RFLP patterns revealed by *Ava*I digests of section 1 internal (S1int) and section 2 (S2) at locus XPL-2. I) *Ava*I digests of section 1 internal (XPL-2S1int). Three patterns are revealed: The first pattern (*A*) is African-specific and is characterized by an 830 bp fragment. The second pattern (*X1*) is common and the third pattern (*X2*) is European-predominant and is characterized by fragments of 600 and 90 bp. II) *Ava*I digests of section 2 (XPL-2S2). For XPL-2R2, three patterns are seen: *X1* and *X2* (common) and *C* (African-specific), characterized by a 1505 bp fragment. The unique combination of patterns revealed in XPL-2S1int and XPL-2S2 is used to determine the allele designation originally found in RFLP analysis of the entire locus. African (AF) and European (EU) alleles are marked with an open and closed line, respectively. Letters in first line are the allele designations initially found in long PCR amplifications, and just above each picture, the new sub-allele designations based on standard PCR amplification.


I
AF
EU
A C D F B E
Long-PCR
S1int
St A X1 X2 X2 X1 X1
Standard-PCR
830
600
90
II
S2
St X1 C X1 X2 X1 X2
1505
700

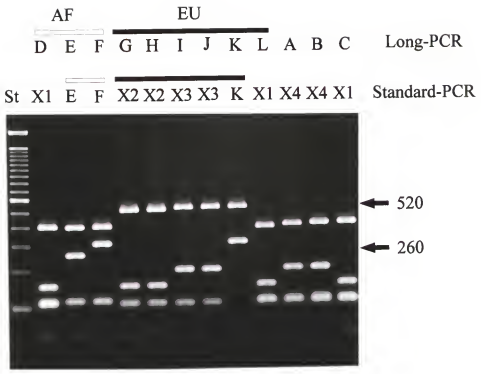
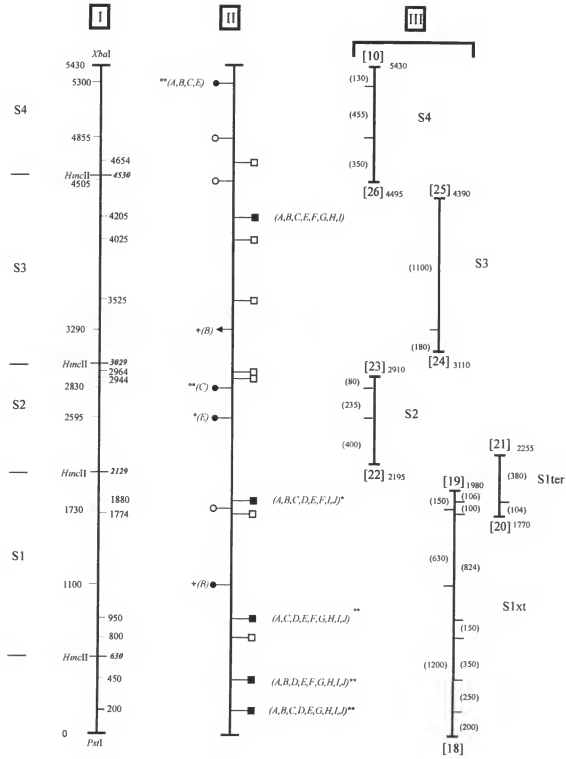


Figure 5-6. RFLP patterns revealed by *Hae*III digests of section 1 internal of locus XPL-2 (XPL-2S1int). African sub-alleles *E* and *F* and the European sub-allele *K* are distinguishable from the others by their unique patterns. Four additional patterns were seen: *X1* and *X4* (common) and *X2* and *X3* (European). European-specific patterns are distinguished by the presence of a 520bp fragment. African-specific sub-allele *E* is characterized by a 260bp fragment. African (AF) and European (EU) alleles are marked with an open and closed line respectively. Letters in the first line are the allele designations initially found in long PCR amplifications, and just above each picture, to the new sub-allele designations based on standard PCR amplification.

Figure 5-7. Restriction cleavage map of locus XPL-5 with the restriction enzymes *Hae*III, *Dde*I and *Spe*I. Numbers and symbols to the left of each map correspond to *Dde*I and *Spe*I sites. Numbers and symbols to the right correspond to *Hae*III sites. Polymorphic and monomorphic sites are indicated by a closed and open symbol respectively. Circles correspond to *Dde*I sites, triangles to *Spe*I sites and squares to *Hae*III sites. Letters in parentheses are the alleles that contain the particular restriction sites. Numbers in brackets are the primers used to amplify each region. Primer sequences are in Table 5-1. Numbers in parentheses are to fragment sizes. I) Relative position of the *Dde*I, *Spe*I and *Hae*III and, the *Hinc*II sites used to subclone locus XPL-5 into several sections amplifiable by standard PCR protocols. II) Relative positions of monomorphic and polymorphic *Dde*I, *Spe*I (left side of the map), and *Hae*III sites (right side of the map). III) Sections of locus XPL-5 amplified with standard PCR protocols. Locus XPL-5 was divided in four sections: section 1 between the two *Hinc*II sites at positions 630 and 2129; section 2 between the *Hinc*II sites at positions 3029 and 2139; section 3 between the *Hinc*II sites at positions 4530 and 3029 and, section 4 between the *Xba*I site at position 5430 and the *Hinc*II site at position 4530. Sites that revealed east European-predominant alleles, west European-predominant alleles, and African alleles are indicated by an asterisk, a (+), and two asterisks, respectively. For *Dde*I, the site at positions 2595 is found in the east European-specific allele *Dde*I-*E,* and the site at position 1100 in the west European-predominant allele *Dde*I-*B*. African-specific alleles are determined by the presence of *Dde*I sites at positions 5300 (*Dde*I-*D*) and 2830 (*Dde*I-*C*). For *Spe*I, the presence of one site at position 3290 results in the west European-predominant allele *Spe*I-*B*. For *Hae*III, three sites located at positions 200, 450 and 950 are responsible for African-specific alleles. The absence of the *Hae*III site located at position 1774 is characteristic of the east European-predominant alleles *Hae*III-*G* and *H*. *Dde*I alleles were determined by amplifying three sections: section 1 extended (S1xt), section 2 (S2) and section 3 (S3). *Spe*I alleles are determined by amplifying one section: section 3 (S3). *Hae*III alleles are determined by amplifying two sections: S1xt and section 1 terminal (S1ter). The polymorphic site that determins the banding pattern for the European predominant alleles *Hae*III-*G* and *H* is included in S1xt. However, the absence of the site at position 1880 in alleles *Hae*III-*G* and *H* results in a 200bp fragment that is indistinguishable from the 200bp fragment generated by the presence of and *Hae*III the site at position 200. Therefore, a small region located at the terminal end of S1xt, and containing the *Hae*III polymorphic site (at position 1880) is amplified and used to determine the presence or absence of this site responsible for the European-predominant alleles *Hae*III-*G* and *H*.


I
II
III
XbaI
5430
5300
S4
4855
4654
HincII
4530
4505
4205
4025
S3
3525
3290
HincII
3029
2964
2944
2830
S2
2595
HincII
2129
1880
1730
1774
S1
1100
950
800
HincII
630
450
200
0
PstI
**(A,B,C,E)
(A,B,C,E,F,G,H,I)
+(B)
**(C)
*(E)
(A,B,C,D,E,F,I,J)*
+(B)
(A,C,D,E,F,G,H,I,J)**
(A,B,D,E,F,G,H,I,J)**
(A,B,C,D,E,G,H,I,J)**
[10] 5430
(130)
(455)
S4
(350)
[26] 4495
[25] 4390
(1100)
S3
(180)
[24] 3110
[23] 2910
(80)
(235)
S2
(400)
[22] 2195
[21] 2255
(380)
S1ter
[19] 1980
(106)
(150)
(100)
(104)
[20] 1770
(630)
(824)
S1xt
(150)
(1200)
(350)
(250)
(200)
[18]

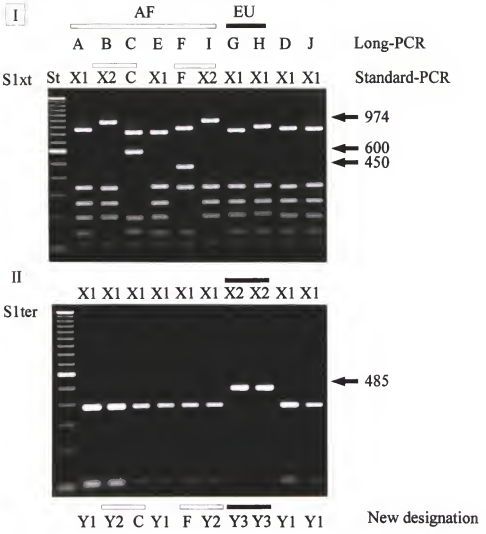
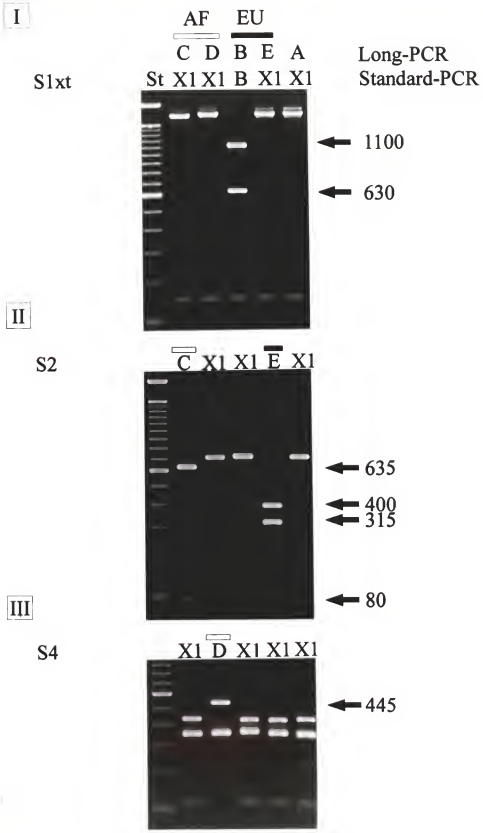


Figure 5-8. RFLP patterns revealed by *Hae*III digests of section 1 extended and section 1 terminal of locus XPL-5. I) Section 1 extended (XPL-5S1xt), and II) section 1 terminal (XPL-5S1ter). Four patterns are seen in digests of XPL-5S1xt: *X1* (common), and *X2*, *C* and *F* (African-specific) which are characterized by fragments of 974 (*X2*), 600 (*C*), and 450 (*F*) bp, respectively. In digests of XPL-5S1ter, two patterns are seen: *X1* (common), and *X2* (European). *X2* is characterized by a 485bp fragment. African (AF) and European (EU) alleles and sub-alleles are marked with an open and closed line, respectively. Letters in first line are to the allele designations initially found in long PCR amplifications, and just above each picture, the new sub-allele designations based on standard PCR amplification relative to each region. Indicated at the bottom are the new sub-allele designations resulting from a combination of RFLP patterns in both sections.

Figure 5-9. RFLP patterns revealed by *Dde*I digests of section1 extended (S1xt), section 2 (S2) and section 4 (S4) of locus XPL-5. I) Patterns revealed in section 1 extended (XPL-5S1xt), II) section 2 (XPL-5S2) and, III) section 4 (XPL-5S4). Four of the initial five alleles (*B*, *C*, *D* and *E*) can be distinguished in separate regions. Allele *Dde*I-A is distinguished by the unique banding pattern combination revealed in all three sections. Allele *Dde*I-*B* is distinguished by RFLP patterns of XPL-5S1xt and is characterized by a 1100 and a 630bp fragment. Alleles *Dde*I-*C* and *E* are distinguished in RFLP patterns of XPL-5S2 and are characterized by fragments of 635 and 80 bp (*Dde*I-*C*) and fragments of 400 and 315 bp (*Dde*I-*E*). Allele *Dde*I-*D* is distinguished by RFLP patterns of XPL-5S4 and is characterized by a 445bp fragment. The common allele, (*Dde*I-*A*) was characterized by the combination of banding patterns in the three sections. African (AF) and European (EU) alleles are marked with an open and closed line respectively. Letters in first line are the allele designations initially found in long PCR amplifications, and just above each picture, the new allele designations, relative to each region, based on standard PCR amplification.


I
AF
EU
C D B E A
Long-PCR
S1xt
St X1 X1 B X1 X1
Standard-PCR
1100
630
II
S2
C X1 X1 E X1
635
400
315
III
80
S4
X1 D X1 X1 X1
445

CHAPTER 6
A MODIFICATION OF THE AFLP PROTOCOL
APPLIED TO HONEY BEE (*APIS MELLIFERA* L.) DNA.

Introduction

Amplified fragment length polymorphism (AFLP) (Vos *et al.*, 1995) is a recently developed PCR-based technique that provides genetic markers for fingerprinting, mapping, and studying genetic relationships among organisms (Becker, *et al.*, 1995; Hill, *et al.*, 1996; Janssen, *et al.*, 1996; Majer, *et al.*, 1996; Thome, *et al.*, 1996; Vos and Kuiper, 1997). This technique is popular because it detects a high amount of polymorphism and is reproducible.

The AFLP protocol developed by Vos *et al,* 1995 involves digesting genomic DNA with two restriction enzymes: a frequent and an infrequent cutter. Following digestion, specially designed adapters for each restriction enzyme are ligated to the ends of the DNA fragments. Primers are made complementary to the adapter sequences, plus the enzyme sequences, and up to three randomly selected nucleotides at the 3' end. The primers are used to amplify the DNA fragments generated by the restriction enzymes in two reactions: primers with one extra base at the 3' end of the enzyme sequences are used in a preamplification, and primers with three extra bases are used in a second amplification. One of the primers in the second reaction is radioactively labeled. The amplified products are resolved in polyacrylamide gels and exposed to film for autoradiography.

Simplification of the original AFLP protocol would facilitate the analysis of large number of samples as, for example, in population genetic studies. Alternative procedures have been proposed to eliminate the use of radioactivity and polyacrylamide gels, but DNA transfer to blots followed by a colorimetric or chemoluminescent detection are still required (Latorra *et al*., 1994; Muller *et al*., 1996). In this report, we present an AFLP protocol simplified and optimized for honey bee (*Apis mellifera* L) DNA. This procedure consists of three steps: 1). Digestion of genomic DNA with a single restriction enzyme and ligation to one type of adapter are accomplished together in one reaction. 2). The digested-ligated DNA is amplified in a single reaction. 3) The AFLP products are resolved in agarose-synergel and visualized by ethidium bromide staining.

Amplification conditions were optimized for concentrations of honey bee template DNA, magnesium ions, primer, and Taq DNA polymerase. This procedure reduces the number of steps and expensive reagents required in the original AFLP protocol. The modified method detects a considerable amount of polymorphism, sufficient for many purposes, but not as much polymorphism as the original procedure. With this modified AFLP procedure, useful honey bee DNA markers were found.

Materials and Methods

DNA Samples.

DNA was isolated from drone honey bees as described in Hall, 1990. Drones are haploid and thus display simpler banding patterns of amplified products than do workers (diploid females). European drone samples came from three locations in the United

States (Florida, Kansas and Tucson), and African samples came from three locations in South Africa (Louis Trichardt, Johannesburg and Pretoria).

Digestion-Ligation Reactions.

*Eco*RI Adapters were prepared by adding equimolar amounts of each strand (Vos *et al*., 1995), and the mix was incubated at 90°C for 1 minute followed by a reduction of 1°C every 20 seconds until room temperature was reached. Digestion of the genomic DNA and ligation with the adapters was done in one reaction. Two µg of genomic DNA were added to 20 µl of 10mM Tris.HAc pH 7.5, 10mM MgAc, 50mM KAc, 5mM dithiotreitol (DTT), 50ng/µl bovine serum albumin (BSA) and 1mM ATP as suggested by Vos *et al*., 1995. Twenty units of *Eco*RI, one unit of T4 DNA ligase and 50 pmol of the *Eco*RI adapter were added and the mixture was incubated at 37°C for five hours. To remove unincorporated adapters and residual enzymes, 80µl of water, 100µ of 7.5M NH_4.Ac and 600µl of 95% ethanol were added. The annealed adapter was precipitated, the pellet washed with 500µl of 75% ethanol, air dried for 10 minutes, resuspended in 100 µl of 10mM Tris.HCl, 0.1mM EDTA (TE) and stored at –20°C. Enzymes were obtained from Boehringer Mannheim (Indianapolis, IN). Primers were synthesized by Lifetechnologies, Gibco BRL(Gaithersburg, MD).

DNA Amplification and Electrophoresis.

As described in the Results section, amplification conditions were optimized for concentrations of four reagents: template honey bee DNA, magnesium ions, primer, and Taq DNA polymerase. PCR reactions were performed in 50µl total volume in thin walled 0.5ml microcentrifuge tubes in an MJ Research PTC-100 thermocycler (Watertown, MA). The PCR profile consisted of one step at 95°C for 1 min followed by 35 cycles of 94°C for 1min, 58°C for 1 min and 72°C for 2 min with a final extension at 72°C for 5min. After optimization of the PCR conditions, African and European honeybee drone DNA samples were amplified with three different primers (*Eco*RI + AGT, *Eco*RI + ATA and *Eco*RI + AGC), and PCR products were resolved in a 2.5% agarose-synergel (0.7% agarose and 0.9% synergel: Diversified Biotech, Boston, MA) in 0.5X TPE buffer or 1X TBE buffer (buffer formulas in Sambrook *et al.,* 1989. Gels were run for at 4V/cm for 14 hours, stained with ethidium bromide, and viewed over UV light to visualize the PCR products.

Results

Optimization of PCR Reagents.

DNA concentration was tested from 1ng to 100ng in a 50µl reaction volume (in triplicate). Above 10ng, high reproducibility with no variability in the banding pattern was seen, but intensification of some bands was apparent as the DNA concentration was

increased. Magnesium ion concentrations were tested from 0.5mM to 3.5mM (in duplicate). At 0.5mM, no bands or a very small number of bands were seen. At 1.5mM, the number of bands increased, and differences in band intensity were seen. Results were reproducible between 2.0 and 2.5mM. As the concentration increased above that level, no new bands were formed, and bands began to disappear until at 3.5mM no bands were seen. Primer concentrations from 100 to 500nM in 100nM increments were tested. Low molecular weight bands were not amplified at 100nM. As primer concentration increased, more bands were visible, and low molecular weight bands increased in intensity. No differences in banding patterns were seen when primers at 400 to 500nM were used. *Taq* DNA polymerase concentration was tested from 0.5 to 3.5 units in a 50µl reaction volume. No amplification was seen at 0.5units. Amplification results were consistent at enzyme concentrations between 1.0 and 3.0 units. At 3.5 units, a small amount of smear was formed. Nucleotide concentrations from 200 to 400µM did not result in differences in the amplification when 2.5mM magnesium ions was used. Thus, the optimized conditions for amplification using honey bee DNA were set at: 2.0mM $MgCl_2$, 2.0 units of *Taq* DNA polymerase, 50ng of DNA, 400nM of primer, and 250µM of each dNTPs in 10mM Tris-HCl pH 8.3, 50mM KCl in a 50µl reaction volume.

Positive and negative controls were used for all amplifications. The positive controls included undigested, unligated DNA with all the other reagents, which checked for any possible bands that could result from primer annealing to a genomic DNA region and not to the adapters. The negative control was a reaction with only the reagents without template DNA and was intended to detect possible artifacts generated from the primers. No visible bands were found in either positive or negative controls, therefore,

all the bands were produced from the annealing of primers to the adapters. Artifactitious bands were seen in the negative control when more than 35 cycles were used in the PCR profile but were not seen in any of the reactions that included template DNA.

AFLP Using Honey Bee DNA.

Three primers were used to amplify DNA of samples from African and European honey bees. The DNA of two drones from each of six colonies, three African and three European, was amplified using the optimized conditions. The amplified products, separated by electrophoresis and stained with ethidium bromide, are seen in Figure 6-1. Substantial amount of polymorphism was seen for all the primers used. A few monomorphic bands (one or two per primer) were produced. Each sample had a unique banding pattern, thereby showing the effectiveness of the method for individual identification (i.e. fingerprinting). For example, with one primer, individuals from different colonies could be distinguished by the presence of unique bands (Figure 6-1-A). African or European-specific bands were also found (Figure 6-1-A; Figure 6-1-B).

Discussion

Population genetic studies have been enhanced with genetic markers detected through methods based on the polymerase chain reaction (PCR), such as microsatellites (Tautz, 1989), restriction fragment length polymorphisms in PCR products (PCR-RFLP) (Karl *et al.*, 1992; Karl and Avise, 1993), random amplified polymorphic DNA (RAPD) (Williams *et al.*, 1990), and more recently, AFLP (Vos *et al.*, 1995). AFLP uses arbitrary

sequences to generate polymorphic amplified fragments from restriction enzyme digested DNA. Thus, this method combines the simplicity of RAPDs with the reliability of RFLPs. As with RAPD analysis, AFLP does not require prior knowledge of the DNA sequence. The longer primers used for AFLP require higher annealing temperature and binding is more specific and stable.

The simplified AFLP procedure described here generates a substantial amount of polymorphism, although, not at levels as high as those generated with the original procedure. The amount of polymorphism can be increased by making primers less specific, (using two extra bases instead of three) by changing the enzyme, (using a four base cutter instead of six) and, by changing the source of DNA (e.g. diploid honey bee workers rather than haploid drones). The digestion of DNA and ligation of adapters is completed in five to eight hours. After the digestion-ligation step, the same sample can be used with different primers, in which case the remaining procedure can be completed in less than 18 hours.

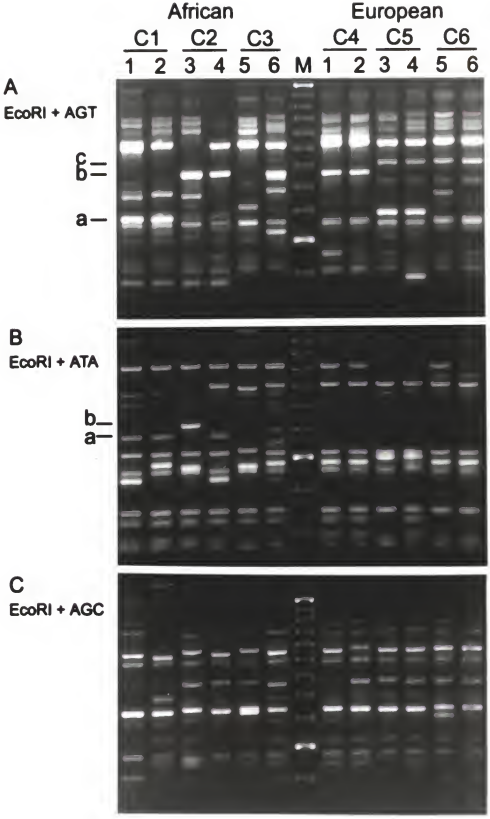
In optimizing the reaction, we found a low sensitivity to the concentrations of DNA and other PCR reagents. High DNA concentrations in AFLP do not affect the reproducibility of the reaction, but the reaction may need to be optimized for DNA from different organisms. In negative controls (all the reagents except DNA), artifactitious bands were detected when more than 35 cycles were used. When genomic DNA was amplified, these bands did not appear even after 35 cycles. Possible problems with AFLP include the coamplification of DNA from endosymbionts such as bacteria or other microorganisms associated with the organism of interest. Partial digestion of the DNA in the digestion-ligation reaction can lead to erroneous results and, therefore, an excess of

restriction enzyme to digest the DNA to completion should be used. Also, comigration of amplified products can make patterns difficult to interpret.

Markers found with AFLP are, in general, dominant. Codominant markers are more valuable for population genetic studies, because homozygous and heterozygous individuals can be distinguished. Codominant markers can be found with AFLP if the polymorphic restriction sites fall within fragments sufficiently short to be amplified. Thus, the fragments lacking the sites and the fragments resulting from cuts at the sites can be seen. By modifying this AFLP procedure to amplify longer DNA fragments (Barnes, 1994; Cheng *et al.*, 1994a and b), the possibility of finding codominant markers would increase.

The AFLP procedure typically displays multiple polymorphisms which can serve as genetic fingerprints of individuals. However, the system can also be used as a source of single genetic markers. DNA fragments containing polymorphic sites can be isolated from the agarose-synergel, cloned, sequenced and used to design primers that amplify only one polymorphic region.

Figure 6-1. Amplification of African and European honey bee drone samples with primers *Eco*RI + AGT (A), *Eco*RI + ATA (B) and *Eco*RI + AGC (C). M= molecular size standard (100bp ladder, Lifetechnologies, Gibco BRL). Two drones per colony (1-6) and three colonies of each bee type (C1-C3 for the African bees and C4-C6 for the European bees) were used. Note the few monomorphic bands and the substantial amount of polymorphism found in the amplification products for all the primers. Bands characteristic of some colonies were found, for example, band "a" for colony 1 and band "b" for colony 2 and 4 using primer *Eco*RI + AGT. African or European-specific bands were also found, for example, band "c" using primer *Eco*RI + AGT is European-specific, and bands "a" and "b" using primer *Eco*RI +ATA are African-specific.


African
European
C1
C2
C3
C4
C5
C6
1 2 3 4 5 6 M 1 2 3 4 5 6
A
EcoRI + AGT
c
b
a
B
EcoRI + ATA
b
a
C
EcoRI + AGC

CHAPTER 7
GENE INTROGRESSION AND HYBRIDIZATION STUDIES OF NEW WORLD HONEY BEE POPULATIONS USING NUCLEAR DNA RFLP MARKERS

Introduction

In the first part of this dissertation, nuclear DNA RFLP markers specific to groups of African, East and West European subspecies were sought. Different methods were employed, including Southern blot analysis with randomly cloned, anonymous, honey bee DNA used as probes, and analysis of anonymous regions amplified by Long PCR. Markers found were made analyzable by standard PCR protocols to facilitate their use. These markers, along with nuclear DNA RFLP markers previously described, were employed to study honey bee populations. The main objectives of the work reported in this chapter were to genetically characterize New World African and European honey bee populations, and to determine levels of hybridization and patterns of gene flow.

Materials and Methods

Honey bee DNA Samples.

Honey bee DNA samples were workers, adults or larvae from Old and New World African and European populations. The Old World African samples were bees of the subspecies *A.m.scutellata* from South Africa. The Old World European samples included subspecies from three major groups: west European (*A.m.mellifera* and

A.m.iberica), east European (*A.m.ligustica* and *A.m.carnica*), and near-eastern (*A.m.caucasica*). Samples from Old World African and European populations have been previously described (See Chapter 2) and were used as references to which New World populations were compared. New World African samples included bees from the Rio Grande Valley in Texas, Mexico, Honduras, Costa Rica, and Venezuela. Samples from Texas were collected from July to September 1995 and from March to May 1997, in what was the northernmost limit of the African bee front in the United States and were provided by O. R. Taylor (University of Kansas, Lawrence, Kansas). African bees were first reported in this area in 1990. Mexican samples were collected in 1989 by H. Glenn Hall (University of Florida, Gainesville, FL) from swarm bait hives in Tapachula, Mexico. Honduran samples were collected in 1989 by ASC from four different locations: El Paraiso, Choluteca, Olancho, and Francisco Morazan. Samples from Costa Rica were collected in 1989 by H. Glenn Hall and H. Arce (Universidad Nacional de Costa Rica) near Cerro de la Muerte and San Isidro del General. Samples from Venezuela were collected in 1986 and 1988 and were provided by R. Hellmich Jr., J. Villa, A. Collins, and T. Rinderer (USDA-ARS, Baton Rouge, Louisiana). Details of number of individuals and colonies used in each of the New World populations are provided in Table 7-1. The New World European samples included bees from a collection of bee stocks across the United States (Page *et al.*, 1982; Severson *et al.*, 1986) and from Kansas (provided by O.R. Taylor, University of Kansas, Lawrence, KS). DNA extraction was performed as described in the Material and Methods section of Chapter 2.

Mitochondrial and Nuclear DNA Markers.

Mitochondrial DNA markers used in this study were RFLPs in three regions. One region was within the cytochrome c oxydase subunit I (CO-I) gene, which contains a *Hinc*II site in west European bees but not in African and east European bees. A second region includes portions of the cytochrome c oxydase subunit I and II (CO-II) gene which has length polymorphisms in African and west European samples. A third region is within the cytochrome b region which contains a *Bgl*II site present in east and west European samples but absent in African samples. Primers and amplification conditions used were as described in Hall and Smith (1991) and in Lee and Hall (1996) without further modification.

Nuclear DNA markers used were those reported earlier in this dissertation (Chapter 2,4 and 5) and summarized in Table 7-2. Amplification conditions, restriction enzyme digestions, and electrophoresis were performed as described in Chapters 2, 4, and 5. In addition to the markers reported here, three loci were used: locus 1231 region E1N digested with *Hha*I (Hall, 1998), a section of locus 130 digested with *Alu*I, and a section of locus 178 digested with *Msp*I , which revealed three African-specific alleles, one east European-specific allele, and one west European-predominant allele, respectively. Primer sequences and amplification conditions for locus 1231 were as described in Hall (1998) without further modification. Loci 130 and 178 were amplified in 50mM Tris-HCl pH 9.2, 16mM $(NH4)_2SO_4$ and 1.5mM $MgCl_2$ with 200µM of each of the four deoxynucleotides (dNTPs), 1 unit of *Taq* DNA polymerase, and 250nM of each primers in a total volume of 25µl. Primers used for 130 were [27] 5'-

TCCGGTACAAAGATAACTAACG-3' and [28] 5'-AACGGAGGTTGTTTGACAC-3'. Primers used for 178 were [29] 5'-TGCGTGATAACGTGGAAATAGCGGA-3' and [30] 5'-ACGCCGAATCTGCCTTTTATTTGCA-3'. Locus 130 was amplified with the following PCR profile: 95°C for 1:00 min, followed by 35 cycles of 94°C for 45sec, 58°C for 45sec, and 72°C for 1:00 min and a final extension at 72°C for 5:00 min. The amplification profile for locus 178 was as for locus 130 but with an annealing temperature of 60°C and an extension of 72°C for 2:00 min. Enzymes and reagents were obtained from Gibco BRL (Lifetechnologies, Gaithersburg, MD).

Allele Counts and Data Analysis.

Genotypes of all the workers for each locus in every population were tabulated. Allele frequencies were calculated from this table using the computer software package GENETIX and GENEPOP version 1.2 (Raymond and Rousset, 1995). Frequency data for loci 1231, 130, and 178 have been already reported in both Old and New World honey bee populations (Hall, 1990; Hall, 1992b; Hall and McMichael, 1996) but were obtained again in this study with the African and European samples from the Old and New World previously described in Chapters 2 and 5.

Fifty-three alleles from eight loci (12 locus enzyme combinations, LECs) were scored, of which seventeen were African-specific, ten were east European-specific and five were west European-specific. Collective frequencies of African, east and west European alleles were calculated for each locus and averaged across loci to obtain mean values for each population (Table 7-2). Markers were tested in 13 Old and New World

populations for a total of 617 individuals from 322 different colonies. Allele frequency data were also used to calculate observed and expected heterozygosity as a measure of genetic diversity as described (Nei 1978; Nei, 1987) using the computer software package GENETIX. Haplotypic diversity values or gene diversity values (H_{div}), which correspond to the proportion of heterozygotes, were calculated as a measure or genetic variation in each population. These values provide the probability of randomly selecting two different alleles in a population and were calculated as described in Nei, (1987) considering all the alleles at all loci examined. Values of H_{div} will vary from "0" (no diversity, population fixed for one allele at each locus, thus only with individuals homozygous at every locus) to "1" (high diversity, populations only with individuals heterozygous at every locus). Exact tests of Hardy-Weinberg equilibrium were performed using the computer software package Arlequin V1.1. Departures from Hardy-Weinberg equilibrium were tested using a modified Markov-chain algorithm (Guo and Thompson, 1992). Table wide significance values were tested and adjusted using a sequential Bonferoni procedure (Holm, 1979; Rice, 1989; Sokal and Rohlf, 1989).

For each bee in every population, the genetic composition was calculated as the proportion of African and European specific alleles across loci. Frequencies of individuals with the same genetic composition were used to determine the degree of hybridization in each population. Hybrid index distributions (Mallet, 1995) were also used to determine levels of hybridization at the population level. African-specific alleles were given a value of "–1" and European alleles "+1". Alleles common to African and European populations were scored with a "0". A score for each allele at each locus was assigned to every worker. Hybrid index values were obtained for each individual by

summing the scores across loci. Frequency distributions of these scores were generated for each population. Mean value comparisons across populations for the different coefficients studied were tested for significance using Duncan multiple range tests (SAS institute). Genetic variation was determined using an analysis of molecular variance (AMOVA), which partitions the total genetic variation into a hierarchical structure (Excoffier *et al.*, 1992) (between groups, among populations within groups, and within populations). AMOVA analysis produced Φ statistics, analogous to Wright's F's statistics, which characterized the variation between groups (Φ_{ct}), between populations within groups (Φ_{sc}) and within populations (Φ_{st}). Six groups were defined for the AMOVA (populations in parentheses): African (*A.m.scutellata*), west European (*A.m.mellifera* and *A.m.iberica*), east European (*A.m.ligustica* and *A.m.carnica*), New World European (United States), New World African (Venezuela, Costa Rica, Honduras, and Mexico) and Texas (Texas 1995 and Texas 1997). Population pairwise Φ_{st} values (or F_{st} values) were calculated as a measure of population differentiation. Significance of Φ statistics and population pairwise F_{st} values were established with permutation tests using 1000 replicates. Gene frequency data were also used for cluster analysis of Old and New World populations. From these frequencies, a genetic distance matrix was calculated using Cavalli-Sforza genetic distance parameter in NTSYS-pc v 1.2 (Exceter software) (Cavalli-Sforza and Edwards, 1967). Cluster analysis of the different populations was performed using the unweighed pair group method with arithmetic average (UPGMA) and by the neighbour-joining method (NJ) (SAHN and NJOIN functions in NTSYS-pc). To evaluate the goodness of fit of the trees produced by the two methods, a cophenetic matrix was calculated from both trees produced by the UPGMA and the NJ method using

the function COPH in NTSYS-pc. These matrices were compared to the genetic distance matrix used to generate the trees. Pearson's correlation coefficients were calculated for each matrix comparison and used as a measure of goodness of fit (Smouse *et al.*, 1986). High values of this correlation coefficient correspond to better fits of the data with the tree produced. Graphical representation of the genetic relationship among populations was also done using a principal coordinate analysis (PCO) with the functions DCENTER and EINGEN in NTSYS.

Results

Mitochondrial DNA Allele Frequencies.

Frequencies for African, east European, and west European mitochondrial DNA in New World African samples are listed in Table 7-3. In the U.S.A. samples, east and west European mitotypes were found in 79 and 21% of the colonies respectively. West European mitotype frequencies in New World European populations were very similar to those found with nuclear DNA markers, reflecting the contribution of the west European subspecies in the North American population (see below). West European mitotypes were found in all the feral samples from the U.S.A. In the Texas 1995 samples, 93 and 7% of the colonies had African and west European mitotypes, respectively. East and west European mitotypes were absent in all the Texas 97 and the New World African populations.

Population Genetic Parameters: Allele Frequencies, Haplotypic Diversity, Hardy-Weinberg Equilibrium

Frequencies for LEC's 1231E1N/*Hha*I, P130/*Alu*I, and P178/*Msp*I in the African subspecies *A.m.scutellata*, the European subspecies *A.m.mellifera*, *A.m.iberica*, *A.m.ligustica*, *A.m.carnica* and *A.m.caucasica,* and the U.S.A. are given in Table 7-4. Allele frequencies for all LECs across New World honey bee populations are shown in Table 7-5.

Collective mean frequencies of African alleles increase in the populations moving toward the south, coinciding with the length of time since African bee invasion. However, no significant differences were found among the collective mean frequencies in the Old and New World African samples including Texas 1995 and Texas 1997 ($P>0.05$) (Table 7-6). African allele frequencies from Texas 1995 were also not significantly different from the U.S.A population. Compared to the Old World African samples, the New World European samples (United States) showed significant differences in the collective mean frequency of African alleles ($P<0.05$).

Collective mean frequencies of east European alleles decreased from 57.9% in the U.S.A populations to 4.2% in the Venezuelan population, which was the most Africanized New World population tested here. For east European alleles, mean frequency values from the U.S.A populations were significantly different from the Old World African populations, and the two most Africanized populations in the New World: Venezuela and Costa Rica ($P<0.05$). Frequencies in the Old World African population were comparable to those in Venezuela and Costa Rica. East European allele frequencies

in the Texas population were comparable to those in Mexico and Honduras and were significantly different from those in Venezuela and Costa Rica.

Collective mean frequencies of west European alleles between New World African and New World European populations were not significantly different ($P>0.05$) but were significantly different from the Old World African population in which west European alleles were found at a mean frequency of 1.5%. Mean frequencies of west European alleles were maintained relatively constant in New World African and European populations.

Haplotypic diversity values for Old and New World African and European populations are shown in Table 7-7. New World African and European populations had significantly higher haplotypic diversity values than the Old World European populations represented by east and west European honey bee subspecies ($P<0.05$) but were not significantly different from the Old World African population. No significant differences were found between the U.S.A. populations and the New World African populations.

No significant departures from Hardy-Weinberg equilibrium were found in any of the Old World European populations. For all the remaining populations, including Old World African, significant departures from Hardy-Weinberg equilibrium were found in 36 tests (23.1%) from a total of 156 possible tests (12 LEC's x 13 populations).

Population Genetic Structure.

Results from partitioning the total genetic variation revealed by AMOVA is presented in Table 7-8. Of the total genetic variation, 69.4% was found within

populations, 23.7% among groups and 6.9% among populations within groups. This high within population variation results in a high level of population differentiation as revealed from the F_{st} value (F_{st}=0.3056 P<0.001). Significant variation among groups (F_{ct}=0.2365 P<0.001) and among populations within groups (F_{sc}=0.09 P<0.001) was also found. Population differentiation as revealed by F_{st} values from pairwise comparisons between populations shows similarities in genetic variation between Texas 1995 and Mexico, and between the populations from Honduras, Costa Rica, and Venezuela (Table 7-9).

Cluster and Principal Coordinate Analyses.

Dendrograms produced from the genetic distance matrix, for all Old and New World honey bee populations, with the UPGMA and NJ methods, are shown in Figure 7-2. Three major groups are clearly distinguishable by both methods: I) An African group formed by all the New World African samples (T95, T97, MEX, HON, CRI and VEN) and the Old World African samples of *A.m.scutellata* (SAF). In both methods, the samples for Texas 1995 are grouped closer with the samples from Mexico. II) A European group formed by the east European samples of *A.m.ligustica* (LIG) and *A.m.carnica* (CAR), by the New World European population (USA), and by the near-east samples of *A.m.caucasica* (CAU). III) A west European group formed by the samples of *A.m.mellifera* (MEL) and *A.m.iberica* (IBE). The west European group is separated and clearly distinguishable from the other European group and the African group in the UPGMA dendrogram but clusters more closely to the African group in the NJ

dendrogram. Despite the similarities found in both UPGMA and NJ methods, the UPGMA clustering provided a better goodness of fit (r=0.88) than the NJ clustering (r=0.71).

Principal coordinate analysis of the genetic similarities of Old and New World populations is presented in Figure 7-3. The first three principal coordinate axes accounted for 72.1%, 33.2% and 8.6% of the total variation. Clusters defined were similar to those found with UPGMA and NJ dendrograms and corresponded to the main groups: west European, east European, the near-east lineage, New World European, New World African, and the Old World African group. Texan populations were grouped with the populations from Honduras and Mexico as in the UPGMA and NJ clustering methods.

Hybridization in New World Honey Bee Populations.

Absolute hybrid index values in all populations studied are listed in Table 7-10, and distributions of the corresponding relative values are shown in Figure 7-4. The Old World African population has a range of hybrid index values from –10 to 0 compared to the east European populations which has a range from 5 to 17. The distribution of hybrid index values in New World European samples is similar to that of the east European samples, whereas the distribution of hybrid index values in New World African samples is closer to the Old World African samples. Samples from the two most Africanized populations in the New World (Venezuela and Costa Rica) have distributions similar to those of the Old World African samples, whereas samples from Texas have distributions

that are intermediate between the Old World African and the New World European samples, suggesting strong hybridization. Moving toward the south, a shift of the hybrid index distributions toward the African distribution is seen in the populations, and most of the intermediate hybrid index values disappear. This trend is also evident in the genotypic composition of workers in the different populations (Figure 7-5) and the number of hybrid workers (Figure 7-6). A dramatic reduction in the frequency and type of hybrids is evident in the populations that had been Africanized for longer periods of time. Frequencies of bees with predominantly African genotypes and/or African and west European genotypes are mainly found in highly Africanized populations in the New World. Although hybrids with east European genotypes are also found, their frequencies were low.

Discussion

Frequency and Distribution of Mitotypes in New World Populations.

The high levels of African mitochondrial DNA found in neotropical populations is consistent with previous studies (Hall and Muralidharan, 1989; Smith *et al.*, 1989). Populations in the subtropical-temperate region in Texas also showed high levels of African mitochondrial DNA, pointing to the continuous expansion of the African bee as unbroken maternal lineages, that is as swarms with African queens (Hall and Muralidharan, 1989).

Nuclear DNA RFLP Allele Frequency Distributions in the New World

Asymmetric gene flow from feral African to managed European neotropical honey bee colonies was reported from previous studies using markers specific to African and east European groups of subspecies (Hall, 1990). European markers were present only at low levels in feral African swarms near the expanding front and were virtually absent behind the front in highly Africanized populations. Managed European colonies behind the front had a greatly reduced level of European alleles, apparently due to considerable African paternal gene flow. In the current study, only feral swarms were tested, and levels of introgression into European colonies were not measured. Therefore, possible asymmetric gene flow was not investigated. However, a more accurate measure of the composition of the feral population was obtained. More samples and populations were tested with a larger collection of markers with broadened specificity.

Earlier studies used east European nDNA markers and did not have nDNA markers specific to west European bees (Hall, 1990). Thus, any west European paternal gene flow could not be detected. In a more recent study, a west European contribution to the feral African population was suggested (McMichael and Hall, 1996). This current study included west European-specific or predominant markers. In populations from tropical and temperate regions in the New World, west European alleles were found at constant frequencies even in highly Africanized populations. The results indicate that west European and African bees do not exist as separate populations. Virtually all west European alleles in the neotropical populations were found associated with African

alleles in individual bees, pointing to admixture between these two subspecies. The original honey bee population in Brazil was largely of the subspecies *A.m.mellifera* (Kent, 1988; Lobo *et al*., 1989; Lobo and Krieger, 1992). The west European alleles in feral neotropical African populations may be due to the predominance of this subspecies in the area of Brazil at the time African bees were introduced. The nearly constant levels of west European markers in feral neotropical African populations suggest that these markers had introgressed early when the African population became established and have persisted as the population expanded.

In European populations of North America (U.S.A.), east European alleles were found at a 70% frequency. West European alleles were found at a 20% frequency. One *A.m.caucasica* specific allele and some African alleles were also found but at very low frequencies. This genetic contribution from different subspecies in the North American population is probably a reflection of previous importations of honey bee subspecies, possibly including non –*scutellata* African subspecies such as *A.m.lamarckii* (Oertel, 1976; Pellet, 1938; Schiff and Sheppard, 1993).

Just south of the expanding African bee front in Texas, all the feral honey bee colonies had African mitochondrial DNA. In these colonies, African nuclear DNA allele frequencies were high, and east European allele frequencies were low. East European allele frequencies increased from south to north. The frequencies of east European alleles and levels of hybridization were highest in the feral African swarms caught in Texas.

These results are consistent with what had been found previously, although generally a higher level of east European alleles were found in neotropical bees than that suggested earlier. The increase in frequency of the east European markers from south to

north may reflect ongoing European paternal introgression (east European drones from apiaries mating with feral African queens) and accumulation of European alleles into the feral African population as it expanded and encountered additional managed European colonies. This introgression could have been enhanced by continuous importation of European queens into commercial apiaries and migratory beekeeping practices, giving the impression that African-east European hybrids persist in the neotropics at higher levels than they actually would under natural conditions. Alternatively, the lower levels of east European alleles in the southern populations may reflect their loss over time as the African bees became established. An apparent loss could be due to continued migration of African bees from more Africanized areas to the south. A loss could be due to selection against hybrids between African and European bees because of poor adaptation to tropical environments or other deficiencies. A loss of European alleles would coincide with changes seen in morphology (Boreham and Roubik, 1987), in allozyme frequencies (Taylor *et al*, 1991), and in DNA allele frequencies between Texas 1995 and Texas 1997 reported in this chapter. As African bees move into an area, frequencies of African alleles should increase and those of European alleles decrease, reaching an equilibrium point (assuming random mating) which would be affected by continuous introduction of European bees, continuous African bee migration from the south, the fitness of hybrids.

As the expanding African bee front reaches its northernmost limit, a hybrid zone may be formed and maintained, as in northern Argentina where African bees have reached their southernmost limit in the Americas (Taylor, 1977; Taylor and Spivak, 1984). Conditions will eventually favor European genotypes as the African bees reach temperate regions, and hybrids may also be better adapted to the environment. The high

degree of hybridization found in Southern Texas and the prevalence of hybrids in a region that has been Africanized for five years is probably an indication that the African migratory front is reaching its limits in North America. However, this has to be determined from further monitoring of the region over a period of several years.

Several mechanisms have been suggested to explain the low levels of fitness of hybrids. These include differences in metabolic capacity, perhaps due to incompatibilities between mitochondrial and nuclear DNA genes (Harrison and Hall, 1993). Selection against European genotypes and hybrids might be the result of reproductive isolating mechanisms that include both behavioral and physiological differences and a better fitness of African bees to tropical environments. Such mechanisms include assortative mating (preferential mating of drones of one type with queens of the same type), queen developmental time (daughter queens with African paternity emerge sooner than those with European paternity), and kin selection (workers raise queens of their own genotype) (DiGrandi Hoffman *et al.*, 1996; Kerr and Bueno, 1970; Page *et al.*, 1989; Winston *et al.*, 1983). Queen developmental time and/or kin selection might select against European genotypes and would be an effective mechanism to eliminate European alleles from the reproductive population that may be expressed in the workers. Hybrids have been found in neotropical conditions, suggesting a lack of selection against them (Rinderer *et al.*, 1991; Sheppard *et al.*, 1991). However, and as explained previously, these hybrids could be the result of constant European gene flow from commercial beekeeping operations into the feral African swarms. In this study, a high level of hybridization has also been found in the subtropical-temperate region in Texas, which would be expected if African bees enter an area densely populated with

commercial European colonies. This however, does not necessarily reflect the adaptability and survival of these hybrids, although higher hybrid fitness is expected in more temperate regions. In highly Africanized populations in the neotropics, hybrids are found in very low frequencies, and most correspond to the African-west European type.

Further DNA studies should help resolve whether or not various processes affect hybrid formation and survival. For example, the availability of multiple markers, more easily analyzable with PCR, enables parental analysis. Genotypes from workers can be used to determine that of their queen mother and the haplotypes of their drone fathers. Hybrid maternity or paternity would suggest hybridization events that happened in previous generations. Hybrid paternal drones would come from hybrid queens, pointing to hybridization events over at least the previous two generations. The absence of parental hybrids and the presence of hybrid workers would suggest recent hybridization. Paternity or maternity not found in expected proportions might point to preferential selection against hybrid drones or queens, thereby pointing to the processes that might be responsible.

The genetic interactions between African and European bees can be measured with gametic phase disequilibria (random association of gametes). Gametic phase disequilibria values approach "0" if populations are randomly mating. The absence of gametic phase disequilibria therefore would suggest admixture between populations, whereas the presence would indicate strong interactions of parental types and possibly selection against hybrids. Gametic phase disequilibria may be an effective measure to detect selection due to factors other than the environment, such as hybrid dysfunction or other isolating mechanisms.

Table 7-1. Number of colonies and individuals used in Old and New World honey bee populations.

Population	No. of individuals	Number of colonies	Number of individuals per colony	Date collected
Old World African				
A.m.scutellata	136	67	2-3	Jan-1990
New World European				
United States	69	44	1-3	
New World African				
Texas 1995	42	14	3	Jul-Sep 1995
Texas 1997	48	16	3	Apr-May 1997
Mexico	38	19	2-3	Jan 1988
Honduras	96	31	2-4	Nov 1989
Costa Rica	23	14	2-3	May 1989
Venezuela	48	19	2-3	Jan-Dec 1988
West European				
A.m.mellifera	13	9	1-2	
A.m.iberica	15	15	1	
East European				
A.m.ligustica	42	41	1-2	
A.m.carnica	26	26	1	
Other European				
A.m.caucasica	21	7	2-3	
Totals	617	322	1-4	

Table 7-2. Nuclear DNA markers used to study hybridization and genetic diversity in New World honey bee populations. Number of alleles per locus and their corresponding frequencies (in parentheses) are shown.

Locus/Enzyme	Total number of alleles	African[a]	East European[b]	West European[c]
XPS-3/*Alu*I	7	4 (0.6288)		1 (1.0000)
XPL-1/*Alu*I	2		1 (0.9497)	
XPL-2/*Ava*I	6	2 (0.5114)	2 (0.8738)	
XPL-2/*Hae*III	7	2 (0.1708)	2 (0.8521)	
XPL-5/*Hae*III	5	3 (0.1148)	1 (0.9547)	
XPL-5/*Dde*I	4	1 (0.2037)	1 (0.3119)	1 (1.0000)
XPL-5/*Spe*I	2			1 (0.5321)
P227-S3xt/*Alu*I	7	2 (0.5280)	1 (0.8960)	
P227-S3xt/*Hinf*I	5		1 (0.6140)	1 (1.0000)
1231E1N/*Hha*I	4	3 (0.6767)		
P130/*Alu*I	2		1 (0.9378)	
P178/*Msp*I	2			1 (0.9121)
Totals[d]	53	17(0.4049)	10(0.7988)	5(0.8888)

(a): Collective frequency of African-specific alleles.
(b): Average collective frequencies in the east European populations *A.m.ligustica* and *A.m.carnica*. Values were calculated by averaging the collective frequencies of east European-specific and east European predominant alleles in both populations.
(c): Average collective frequencies in the west European populations *A.m.mellifera* and *A.m.iberica*. Values were calculated as in the east European populations.
(d): Values in parentheses correspond to the averages across loci.

Table 7-3. Frequencies of mitochondrial DNA types in New World honey bee populations

Population	Number of colonies	Mitotype		
		African	East European	West European
United States	44	0	0.79	0.21
Texas 1995	14	0.93		0.07
Texas 1997	16	1.00		
Mexico	19	1.00		
Honduras	31	1.00		
Costa Rica	14	1.00		
Venezuela	19	1.00		

Table 7-4: Collective African (C_{AF}), east European (C_{EE}) and west European (C_{WE}) frequencies for alleles at loci 1231E1N, 130 and 178 in African and European populations and observed and expected heterozygosity (H_{obs} and H_{exp} respectively).

Locus/Enzyme Allele	Populations						
	African		West European		East European		Other
	South Africa (*A.m.scutellata*)	United States	France (*A.m.mellifera*)	Spain (*A.m.iberica*)	Italy (*A.m.ligustica*)	Austria (*A.m.carnica*)	Rusia (*A.m.caucasica*)
1231E1N/*Hha*I							
(N)	133	68	13	15	42	26	20
X2[¥]	0.1316						
X3[¥]	0.3233			0.0333			
X4[¥]	0.2218						
X1	0.3233	1.0000	1.0000	0.9667	1.0000	1.0000	1.0000
C_{AF}	0.6766			0.0333			
H_{obs}	0.7068	0.0000	0.0000	0.0667	0.0000	0.0000	0.0000
H_{exp}	0.7244	0.0000	0.0000	0.0667	0.0000	0.0000	0.0000
130/*Alu*I							
(N)	136	69	13	15	42	26	20
E[§]	0.0037	0.6812	0.0769		0.9524	0.9231	
O	0.9963	0.3188	0.9231	1.0000	0.0476	0.0769	1.0000
C_{EE}	0.0037	0.6812	0.0769		0.9524	0.9231	
H_{obs}	0.0074	0.5217	0.1538	0.0000	0.0952	0.1538	0.0000
H_{exp}	0.0073	0.4344	0.1420	0.0000	0.0907	0.1420	0.0000

Table 7-4 continued

178/*Msp*I

(N)	136	66	12	15	41	26	20
A	0.9449	0.8333	0.0417	0.1333	0.9512	0.9423	0.9250
$B^{\S\S}$	0.0551	0.1667	0.9583	0.8667	0.0488	0.0577	0.0750
C_{WE}	0.0551	0.1667	0.9583	0.8667	0.0488	0.0577	0.0750
H_{obs}	0.0809	0.3333	0.0833	0.2667	0.0976	0.1154	0.1500
H_{exp}	0.1042	0.2778	0.0799	0.2311	0.0928	0.1087	0.1387

	South Africa (*A.m.scutellata*)	United States	France (*A.m.mellifera*)	Spain (*A.m.iberica*)	Italy (*A.m.ligustica*)	Austria (*A.m.carnica*)	France (*A.m.caucasica*)
Average number of alleles/locus(*)	3.4167	2.8333	1.5000	1.3333	2.0833	2.4167	1.75
H_{obs} averaged across loci(*)	0.2927 ±0.2689	0.3929 ±0.1973	0.0972 ±0.1666	0.1218 ±0.2246	0.2074 ±0.2683	0.2012 ±0.1601	0.1590 ±0.2151
H_{exp} averaged across loci(*)	0.3452 ±0.3141	0.4356 ±0.1880	0.0989 ±0.1670	0.0972 ±0.1714	0.1713 ±0.1815	0.2251 ±0.1809	0.1526 ±0.2087

(*): Values include data from alleles of locus P227-S3xt, XPL-1, XPL-2, XPL-5, and XPS-3.

(§): East European-specific or predominant allele or sub-allele.

(§§): West European-predominant allele or sub-allele.

(¥): African-specific or predominant allele or sub-allele.

(*): Alleles found in very low frequencies and were not detected in worker samples but in one drone sample.

(N): Number of workers (2N= number of alleles).

Table 7-5. Collective African (C_{AF}), east European (C_{EE}) and west European (C_{WE}) allele frequencies in New World honey bee populations and observed and expected heterozygosity (H_{obs} and H_{exp} respectively).

Locus/Enzyme Allele	Population					
	Texas 1995	Texas 1997	Mexico	Honduras	Costa Rica	Venezuela
XPL-1/*Alu*I						
(N)	47	43	38	96	23	46
G						
X1	0.7872	0.7326	0.8289	0.8438	0.8913	0.9239
X2[§]	0.2128	0.2674	0.1711	0.1563	0.1087	0.0761
C_{EE}	0.2128	0.2674	0.1711	0.1563	0.1087	0.0761
H_{obs}	0.2979	0.2558	0.3421	0.2292	0.2174	0.1087
H_{exp}	0.3350	0.3918	0.2836	0.2637	0.1938	0.1406
XPL-2/*Ava*I						
(N)	45	40	38	80	23	44
A[*]	0.3111	0.2125	0.2763	0.2438	0.3261	0.2386
C[*]	0.0778	0.1375	0.0921	0.1500	0.0652	0.0795
D[§]	0.2111	0.1500	0.2337	0.1250	0.1087	0.0568
F[§]					0.0217	
B	0.4	0.4625	0.3816	0.4812	0.4565	0.6136
E		0.0375	0.0263		0.0217	0.0114
C_{AF}	0.3889	0.3500	0.3684	0.3938	0.3913	0.3181
C_{EE}	0.2111	0.1500	0.2337	0.1250	0.1299	0.0568
H_{obs}	0.6667	0.7500	0.7368	0.7000	0.7826	0.5455
H_{exp}	0.6926	0.6981	0.7188	0.6709	0.6682	0.5568
XPL-2/*Hae*III						
(N)	46	40	36	94	22	42

Table 7-5 continued.

E[*]	0.1087	0.0625	0.0278	0.0372		0.0833
F[*]	0.0761		0.0417	0.0053	0.0909	
X2[§]	0.1196	0.1250	0.1111	0.0638	0.1136	0.0357
X3[§]	0.0326	0.0500	0.1250	0.0160		0.0119
K						
X1	0.2500	0.3750	0.3611	0.3883	0.3409	0.4048
X4	0.4130	0.3875	0.3333	0.4894	0.4545	0.4643
C_{AF}	0.1848	0.0625	0.0695	0.0425	0.0909	0.0833
C_{EE}	0.1522	0.1750	0.2361	0.0798	0.1136	0.0476
H_{obs}	0.6739	0.6250	0.6389	0.5213	0.5909	0.3810
H_{exp}	0.7339	0.6872	0.7280	0.6040	0.6560	0.6122
XPL-5/*Hae*III						
(N)	45	39	38	87	23	41
Y2[*]				0.0115		
C[**]						
F	0.0111				0.0217	0.0122
Y4	0.0222	0.0897	0.0263	0.0575	0.0435	0.0244
Y3[§]	0.3222	0.3462	0.3289	0.0805	0.0652	0.0122
Y1	0.6444	0.5641	0.6447	0.8506	0.8696	0.9512
C_{AF}				0.0115		
C_{EE}	0.3222	0.3462	0.3289	0.0805	0.0652	0.0122
H_{obs}	0.3556	0.2051	0.3421	0.1609	0.2609	0.0976
H_{exp}	0.4802	0.5539	0.4754	0.2666	0.2372	0.0943
XPL-5/*Dde*I						
(N)	47	42	38	93	22	46
C[*]						

Table 7-5 continued.

D[*]	0.0957	0.0952	0.1316	0.1452	0.1136	0.1304
B[§]	0.2340	0.2381	0.2632	0.2366	0.2273	0.1957
E[§]	0.0106	0.0595	0.0263		0.0227	
A	0.6596	0.6071	0.5789	0.6183	0.6364	0.6739
C_{AF}	0.0957	0.0952	0.1316	0.1452	0.1136	0.1304
C_{WE}	0.2340	0.2381	0.2632	0.2366	0.2273	0.1957
C_{EE}	0.0106	0.0595	0.0263		0.0227	
H_{obs}	0.5532	0.5476	0.6316	0.5161	0.4091	0.4783
H_{exp}	0.5009	0.5621	0.5776	0.5407	0.5300	0.4905
XPL-5/*Spe*I						
(N)	47	43	38	93	23	40
A	0.9043	0.8721	0.9079	0.8441	0.9348	0.9375
B[§§]	0.0957	0.1279	0.0921	0.1559	0.0652	0.0625
C_{WE}	0.0957	0.1279	0.0921	0.1559	0.0652	0.0625
H_{obs}	0.1915	0.2093	0.1842	0.1613	0.1304	0.1250
H_{exp}	0.1732	0.2331	0.1672	0.2632	0.1219	0.1172
P227-S3xt/*Alu*I						
(N)	45	42	37	86	23	30
A[*]	0.2444	0.2381	0.2838	0.3256	0.4535	0.3333
A1	0.0000					
J[*]	0.3222	0.4286	0.4324	0.4128	0.4783	0.5167
I[§]	0.4111	0.2619	0.2297	0.2500	0.0652	0.1333
E[§]	0.0000	0.0595	0.0541	0.0116		
F	0.0111	0.0119				0.0167
B	0.0111					

Table 7-5 continued

C_{AF}	0.5666	0.6667	0.7162	0.7384	0.9318	0.8500
C_{EE}	0.4111	0.3214	0.2838	0.2616	0.0652	0.1333
H_{obs}	0.2667	0.4762	0.4595	0.5116	0.3913	0.3667
H_{exp}	0.6672	0.6874	0.6768	0.6610	0.5586	0.6039
P227-S3xt/*Hinf*I						
(N)	45	42	37	85	23	29
$F^{\S}$	0.1111	0.0357	0.1486	0.1471		
$G^{\S\S}$	0.3222	0.3214	0.0811	0.1941	0.2174	0.2759
$H^{\S}$	0.0222			0.0059		
D	0.4778	0.5833	0.7162	0.6059	0.7391	0.7241
$Y1^{**}$	0.0667	0.0595	0.0541	0.0471	0.0435	
C_{EE}	0.1333	0.0357	0.1486	0.1530		
C_{WE}	0.3222	0.3214	0.0811	0.1941	0.2174	0.2759
H_{obs}	0.2667	0.1667	0.1892	0.2118	0.0000	0.0000
H_{exp}	0.6506	0.5516	0.4554	0.5713	0.4045	0.3995
P178/*Msp*I						
(N)	45	38	38	89	22	46
A	0.8000	0.8421	0.7763	0.8652	0.8636	0.8261
$B^{\S\S}$	0.2000	0.1579	0.2237	0.1348	0.1364	0.1739
C_{WE}	0.2000	0.1579	0.2237	0.1348	0.1364	0.1739
H_{obs}	0.3556	0.1579	0.1842	0.2472	0.1818	0.3043
H_{exp}	0.3200	0.2659	0.3473	0.2333	0.2355	0.2873
1231/HhaI						
(N)	46	43	38	87	23	45
$X2^{*}$	0.1957	0.0930	0.1447	0.0862	0.0435	0.1111

Table 7-5 continued

X3[¥]	0.1413	0.1047	0.1053	0.2241	0.2826	0.1889
X4[¥]	0.1522	0.0581	0.0921	0.1149	0.1522	0.1667
X1	0.5109	0.7442	0.6579	0.5747	0.5217	0.5333
C_{AF}	0.4892	0.2558	0.3421	0.4252	0.4783	0.4667
H_{obs}	0.6739	0.2791	0.3684	0.5632	0.4783	0.6444
H_{exp}	0.6576	0.4232	0.5267	0.5988	0.6229	0.6398
P130/*Alu*I						
(N)	47	41	38	91	23	46
E[$]	0.2128	0.1951	0.1974	0.1264	0.0435	0.0109
O	0.7872	0.8049	0.8026	0.8736	0.9565	0.9891
C_{EE}	0.2128	0.1951	0.1974	0.1264	0.0435	0.0109
H_{obs}	0.3830	0.2927	0.3421	0.1648	0.0870	0.0217
H_{exp}	0.3350	0.3141	0.3168	0.2208	0.0832	0.0215
XPS-3/*Alu*I						
(N)	47	42	38	84	22	43
A[¥]	0.2872	0.1310	0.1184	0.2321	0.2955	0.2907
B[¥]	0.2340	0.1667	0.2763	0.1726	0.2045	0.2442
C[¥]				0.0060		
D[¥]	0.0532	0.0595	0.0658	0.2024	0.0227	0.0349
E		0.0119		0.0119		
G	0.3830	0.4286	0.3026	0.3393	0.3182	0.2907
H[§§]	0.0426	0.2024	0.2368	0.0357	0.1591	0.1395

Table 7-5 continued

C_{AF}	0.5744	0.3572	0.4605	0.6131	0.5227	0.5698
C_{WE}	0.0426	0.2024	0.2368	0.0357	0.1591	0.1395
H_{obs}	0.7234	0.6429	0.7368	0.7500	0.5455	0.6047
H_{exp}	0.7114	0.7268	0.7576	0.7588	0.7438	0.7507

	Texas 1995	Texas 1997	Mexico	Honduras	Costa Rica	Venezuela
Average number Of alleles/locus	3.7500	3.6667	3.5833	3.7500	3.4167	3.3333
H_{obs} averaged across loci.	0.4507 ±0.1936	0.3840 ±0.2113	0.4297 ±0.2087	0.3948 ±0.2209	0.3396 ±0.2331	0.3065 ±0.2315
H_{exp} averaged Across loci	0.5215 ±0.1908	0.5071 ±0.1792	0.5026 ±0.1957	0.4711 ±0.2033	0.4213 ±0.2366	0.3923 ±0.2515

(§): East European-specific or predominant allele or sub-allele.
(§§): West European-predominant allele or sub-allele.
(¥): African-specific or predominant allele or sub-allele.
(*): Alleles found in very low frequencies and were not detected in worker samples but in one drone sample.
(**): New allele.
(N): Number of workers (2N= number of alleles).

Table 7-6. Average collective frequencies of African, east and west European alleles in New World and South African honey bee populations. Means with the same letters are not significantly different.

	Allele specificity[1]		
Population	African (n=7)	East (n=8)	West (n=5)
United States	0.006 ± 0.000 b	0.579 ± 0.034 a	0.276 ± 0.006 a
Texas 1995	0.255 ± 0.045 a b	0.194 ± 0.011 a	0.210 ± 0.005 a
Texas 1997	0.329 ± 0.047 a	0.208 ± 0.013 a	0.179 ± 0.010 a
Mexico	0.298 ± 0.054 a	0.203 ± 0.007 a	0.179 ± 0.006 a
Honduras	0.339 ± 0.068 a	0.123 ± 0.005 ac	0.151 ± 0.005 a
Costa Rica	0.361 ± 0.090 a	0.069 ± 0.002 cd	0.161 ± 0.003 a
Venezuela	0.345 ± 0.079 a	0.042 ± 0.002 cd	0.170 ± 0.005 a
South Africa	0.403 ± 0.048 a	0.012 ± 0.000 d	0.016 ± 0.000 b
Average	0.292 ± 0.054	0.179 ± 0.009	0.168 ± 0.004

(1): n=number of locus enzyme combinations (LEC's)

Table 7-7. Haplotypic diversity values in Old and New World honey bee populations. Means with the same letter are not significantly different.

Population	n[(1)]	Haplotypic diversity (H_{div}) ± s.e
Costa Rica	12	0.5296 ± 0.0313 a
Texas 1995	12	0.5279 ± 0.0376 a
Texas 1997	12	0.5136 ± 0.0330 a
Mexico	12	0.5097 ± 0.0396 a
Honduras	12	0.4722 ± 0.0427 a
United States	12	0.4396 ± 0.0359 a
Venezuela	12	0.4101 ± 0.0706 a
A.m.scutellata	12	0.3484 ± 0.0988 ad
A.m.carnica	12	0.2296 ± 0.0340 cd
A.m.ligustica	12	0.1735 ± 0.0337 cd
A.m.caucasica	12	0.1590 ± 0.0450 c
A.m.mellifera	12	0.1020 ± 0.0293 c
A.m.iberica	12	0.1006 ± 0.0315 c

(1): n=number of locus enzyme combinations (LEC's).

Table 7-8. AMOVA results. P-values determined from permutation tests using 1000 replicates.

Source of variation	Φ statistic	F_{st} statistic	% of variation	P-value
Among groups	0.71373	F_{ct}=0.23647	23.65	<0.0001
Among populations within groups	0.20876	F_{sc}=0.09059	6.92	<0.0001
Within populations.	2.09573	F_{st}=0.30564	69.44	0.00293

Table 7-9. Pairwise F_{st} values for Old and New World honey bee populations and corresponding significance levels.

	SAF	MEL	IBE	LIG	CAR	CAU	USA	T95	T97	MEX	HON	CRI
MEL	0.5550*											
IBE	0.5183*	0.1818*										
LIG	0.5727*	0.8072*	0.8086*									
CAR	0.5692*	0.7646*	0.7663*	0.1197*								
CAU	0.3434*	0.7838*	0.7690*	0.5865*	0.5823*							
USA	0.4375*	0.4880*	0.4985*	0.1372*	0.0927*	0.3211*						
T95	0.1156*	0.3775*	0.3616*	0.3944*	0.3615*	0.2480*	0.2167*					
T97	0.1342*	0.3813*	0.3652*	0.4014*	0.3656*	0.2178*	0.1999*	0.0098*				
MEX	0.0998*	0.3942*	0.3756*	0.4244*	0.3871*	0.2305*	0.2275*	**0.0027ns**	0.0133*			
HON	0.0770*	0.4067*	0.3860*	0.4619*	0.4432*	0.2868*	0.2878*	0.0223*	0.0154*	0.0149*		
CRI	0.0571*	0.4786*	0.4474*	0.5756	0.5290*	0.3379*	0.3409*	0.0391*	0.0356*	0.0223*	**0.0093ns**	
VEN	0.0523*	0.4707*	0.4382*	0.5745*	0.5467*	0.3486*	0.3516*	0.0273	0.0193*	0.0186*	**-0.0018ns**	**-0.0203ns**

(*): P<0.001.

ns: Not significant values, also indicated in bold and underlined numbers.

Populations are: SAF= *A.m.scutellata* (South Africa), MEL=*A.m.mellifera* (France), IBE=*A.m.iberica* (Spain), LIG=*A.m.ligustica* (Italy), CAR=*A.m.carnica* (Austria), CAU=*A.m.caucasica* (Rusia), USA=United States, T95=Texas 1995, T97=Texas 1997, MEX=Mexico, HON=Honduras, CRI=Costa Rica and VEN=Venezuela.

Table 7-10. Absolute frequencies of hybrid index values for Old and New World honey bee populations.

			Hybrid index value																																
Pop	M	S	-10	-	9	-	8	-	7	-	6	-	5	-	4	-	3	-	2	-	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
MEL	0.77	1.30																	1		1	3	3	5														
IBE	-0.73	1.22																	6		1	7		1														
LIG	11.76	1.68																												1	4	4	10	6	11	5	1	
CAR	12.04	3.31																													3	3	4	3	9	2	1	1
CAU	7.14	2.03																					1				1	3	7	3	6							
SAF	-5.93	1.54					12		35		37		22		17		5		2		1																
USA	9.38	2.62																				1			1	2		3	5	14	9	8	10	9	6	1	
T95	-0.79	3.55									1		4		7		3		5		3	6	6	1		3	1				2						
T97	-1.21	3.06									1		4		6		5		7		6	6	2	4	3	1	1		1								
MEX	3.95	5.89			1		2		1		3		4		3		1		2		2	5	1	5	1	2	2	1	2								
HON	-2.59	2.90			2		2		2		5		13		17		16		6		9	11	5	3	3		1	1									
CRI	-4.09	2.67					1		2		5		2		5		2		4			1				1											
VEN	-3.48	2.40	1						4		5		6		8		8		4		7	3	2														

Pop=populations: MEL=*A.m.mellifera*, IBE=*A.m.iberica*, LIG=*A.m.ligustica*, CAR=*A.m.carnica*, CAU=*A.m.caucasica*, SAF=*A.m.scutellata* , USA=United States, T95=Texas 1995, T97=Texas 1997, MEX=Mexico, HON=Honduras, CRI=Costa Rica, VEN=Venezuela. M=Mean and S=Standard error.

African alleles were scored with "–1", European alleles with "+1" and common alleles with "0".

M=Mean hybrid index value.

S=Standard error of the mean hybrid index value.

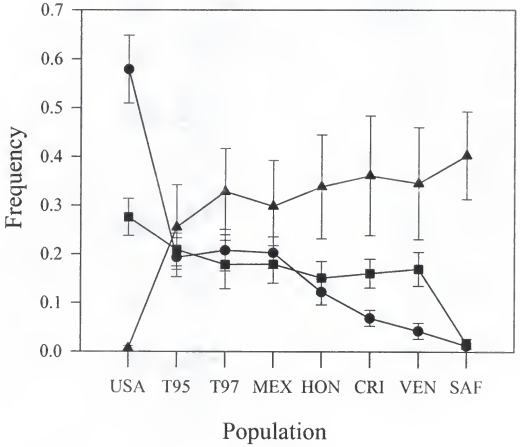
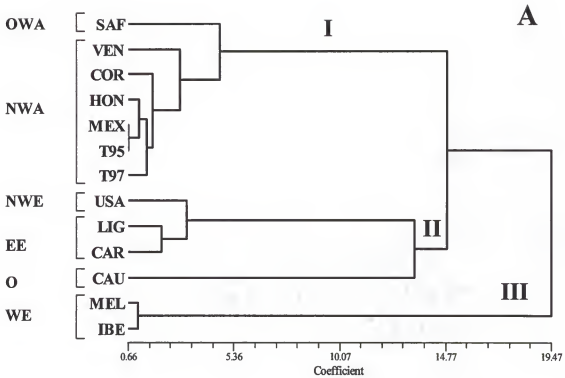
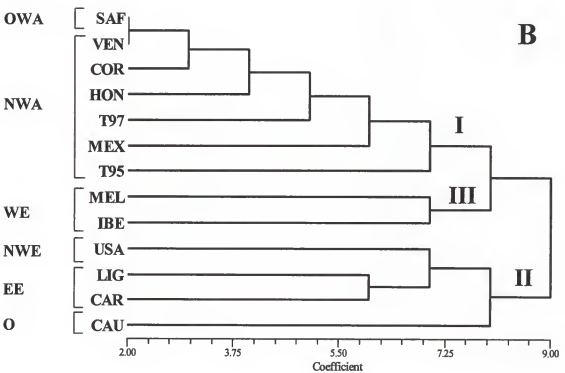


Figure 7-1. Collective mean frequencies of African (▲) East (●) and west (■) European alleles in New World African and European populations and in Old World African populations.

Figure 7-2. Cluster analyses of Old and New World honey bee populations. A) Analysis using the unweighted paired group mean arithmetic (UPGMA) and B) analysis using the neighbour joining method (NJ). Trees were constructed using the Cavalli-Sforza genetic distance coefficient. OWA=Old World African, NWA=New World African, NWE=New World European, EE=east European, WE=west European and O=Other (near-east lineage: *A.m.caucasica*). Groups represented are: African (I), east and near-east Europeans (II) and west European (III).


A
OWA
NWA
NWE
EE
O
WE
SAF
VEN
COR
HON
MEX
T95
T97
USA
LIG
CAR
CAU
MEL
IBE
I
II
III
0.66
5.36
10.07
14.77
19.47
Coefficient


B
OWA
NWA
WE
NWE
EE
O
SAF
VEN
COR
HON
T97
MEX
T95
MEL
IBE
USA
LIG
CAR
CAU
I
III
II
2.00
3.75
5.50
7.25
9.00
Coefficient

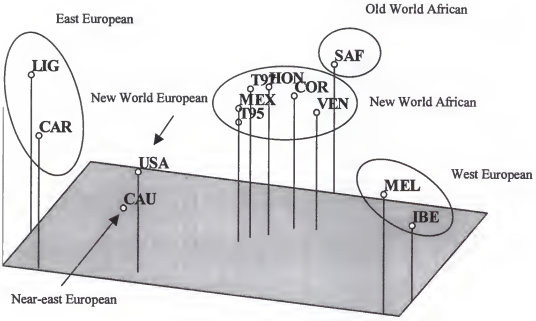


Figure 7-3. Three dimensional representation of principal coordinate analysis of New and Old World honey bee populations as inferred from the genetic distance matrix using the Cavalli-Sforza distance parameter. Population abbreviations are as in Table 7-9.

Figure 7-4. Hybrid index distributions. A) *A.m.scutellata* (dark bars) and United States (light bars), B) Texas 1995, C) Texas 1997, D) Mexico (light bars) and Honduras (dark bars), E) Costa Rica (light bars) and Venezuela (dark bars), and F) Old World European: *A.m.mellifera* (hatched bars), *A.m.iberica* (open bars), *A.m.ligustica* (crossed bars), *A.m.carnica* (dark bars) and *A.m.caucasica* (light bars).

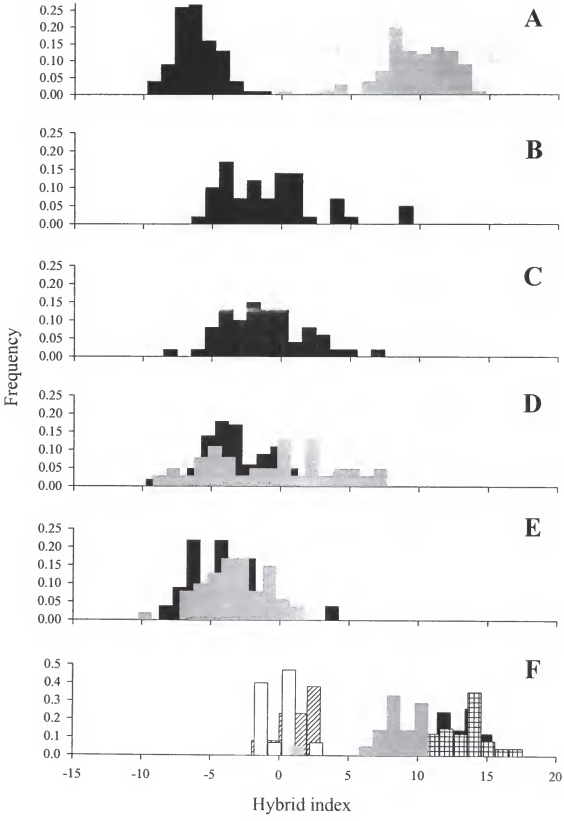
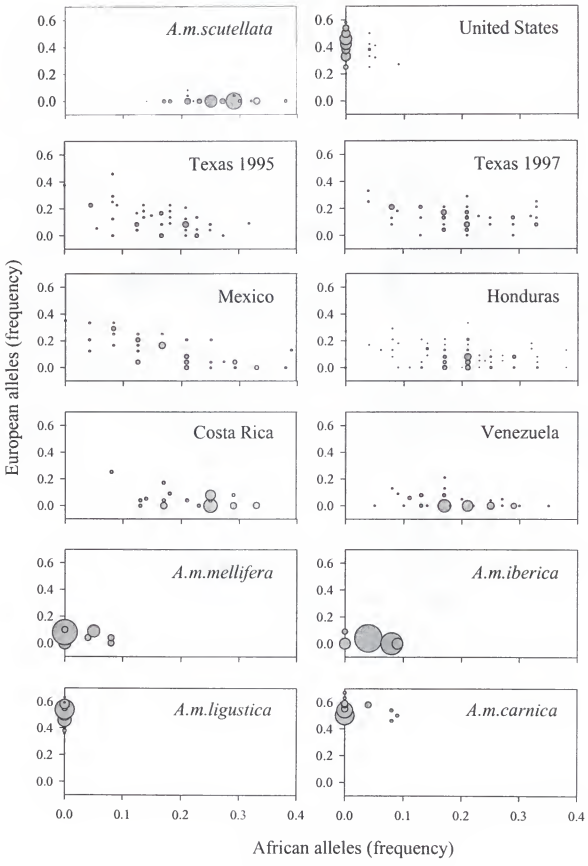

A
B
C
D
E
F
Frequency
Hybrid index
0.25
0.20
0.15
0.10
0.05
0.00
0.5
0.4
0.3
0.2
0.1
0.0
-15
-10
-5
0
5
10
15
20

Figure 7-5. Bubble plots representing the genotypic composition of Old and New World honey bee populations. Frequencies of workers with a specific genotypic composition are in proportion to the size of the bubble. New World African populations (Texas 95 to Venezuela) are arranged according to a north to south geographical location coinciding with increasing times since the African bees invasion.


A.m.scutellata
United States
Texas 1995
Texas 1997
Mexico
Honduras
Costa Rica
Venezuela
A.m.mellifera
A.m.iberica
A.m.ligustica
A.m.carnica
European alleles (frequency)
African alleles (frequency)
0.0
0.2
0.4
0.6
0.1
0.3

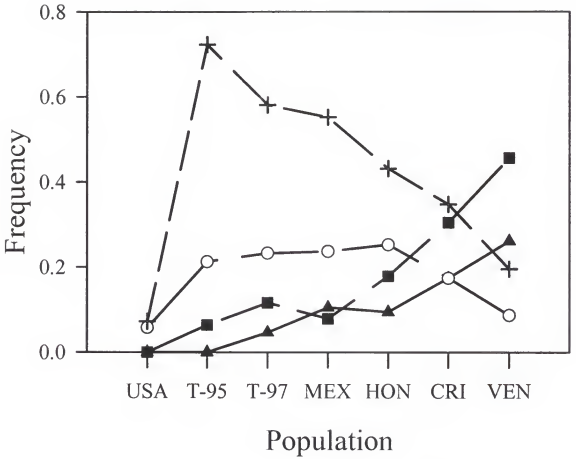


Figure 7-6. Frequencies of African-European hybrid workers in New World honey bee populations. African (▲), African-east European (O), African-west European (■) and, African-east-west European (✣) hybrid distributions correspond to the proportion of workers with a corresponding hybrid type relative to the nuclear DNA genotypes.

CHAPTER 8
CONCLUDING REMARKS

The arrival of the African honey bee to the United States and its dispersal to most of the southern states will have a negative impact in the beekeeping industry in this country. Behavioral characteristics including high swarming, absconding, and stinging make African bees difficult to manage (Michener *et al.*, 1972; Michener, 1975). Commercial beekeeping operations with African bees will require different methods of handling and managing colonies and will be subject to restrictions on locations and movement of beehives. Alternatively, beekeepers will have to rely on certified, desirable breeding stocks. Although income from honey production amounts to only 150 million dollars a year, the most important value of beekeeping in the United States is in pollination, which directly and indirectly represent benefits estimated at 10 to 20 billion dollars a year (McDowell, 1984; Levin, 1983).

Accurate identification of African and European bees is necessary for research, regulation and commercial beekeeping. Prior identification has been limited to morphometrics, cuticular hydrocarbons, allozymes, and mitochondrial DNA. These characteristics, however, have their limitations as was explained in previous chapters. With previous genetic markers, hybridization studies have been limited due to the low levels of polymorphisms detected (allozymes) and the uniparental mode of inheritance (mitochondrial DNA). To understand better the processes involved in the spread of the African bees and the genetic interactions between African and European subspecies,

nuclear genetic markers with high levels of polymorphism, that are biparentally inherited, are needed. Nuclear DNA RFLP markers specific to different subspecies of honey bees including east European (*A.m.ligustica* and *A.m.carnica*), west European (*A.m.iberica* and *A.m.mellifera*), near-eastern (*A.m.caucasica*) and African (*A.m.scutellata*) have been sought using different methodologies. RFLP markers were initially found using laborious procedures, that is with Southern blots and randomly cloned honey bee DNA probes. RFLP markers first found with probes were made analyzable by standard PCR procedures (Hall, 1998), which do not require high quality and quantities of DNA. Thus, large number of samples could be tested at a relatively low cost.

As reported in later chapters, markers were found in DNA first amplified with PCR. To enhance this search, markers were found in DNA amplified by long PCR and were subsequently made analyzable by standard PCR. The nuclear DNA PCR-RFLP markers specific to groups of honey bee subspecies described in this dissertation, make for an important addition to those that already exist or that are still being converted to a PCR analyzable format. Despite the collection of markers now available, more markers specific to African and west European bees will be needed in the future. Most of the markers now available are specific for the subspecies groups, but no single allele is diagnostic (Ayala and Powell, 1972) that is, specific and nearly fixed in a population. African and west European-specific markers are difficult to find due to the high genetic diversity in the African population, and to the effects of genetic pollution that result from migratory beekeeping practices and importation of queens from other environments.

The markers specific to different honey bee subspecies were successfully employed in this research to characterize the subspecies composition and the extent of

hybridization in New World populations including a region in southern Texas close to the predicted hybrid zone. Evidence of a high degree of Africanization in New World populations was obtained, however, a high degree of hybridization was also apparent during the first years after African bee invasion. An apparent lost of east European alleles was seen, greater in populations farther south that had been Africanized for longer periods of time, suggesting selection against bees with east European genes. West European allele frequencies however, remained relatively constant, suggesting the possibility of an admixed African-west European population.

These markers will continue to be valuable for ongoing studies in the neotropics, for studies of the hybrid zone expected to form in the southern portion of the United States, and for studies of African gene introgression north of the hybrid zone where European bees are predominantly found. Future research results may point to mechanisms that might isolate the African and European subspecies and may reveal if the apparent loss of hybrids in the neotropics reflects poor adaptation to the environment or if other hybrid deficiencies are involved. Whatever the cause, hybrid survival is an important factor to consider in hybridization programs proposed to minimize the impact of Africanization in the United States, for example, genetic dilution of feral African bees or the development of African-European hybrids for commercial beekeeping. Also, such studies of African-European interactions should provide a better understanding of the evolutionary processes involved in subspecies differentiation resulting from their geographical distribution.

REFERENCES

Adams, J., Rothman, E.D., Kerr, W.E. and Paulino, Z.L. 1977. Estimation of the number of sex alleles and queen matings from diploid male frequencies in a population of *Apis mellifera*. *Genetics* 86: 583-596.

Anglani, F., Picci, L., Camporese, Ch. and Zacchello, F. 1990. Heteroduplex formation in Polymerase Chain Reaction. *Am. J. Hum. Genet.* 47: 169-170.

Arias, M.C. and Sheppard, W.S. 1996. Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) inferred from mitochondrial DNA sequence. *Mol. Phylogenet. Evol.* 5(3): 557-566.

Arnheim, N. and Erlich, H. 1992. Polymerase Chain Reaction strategy. *Ann. Rev. Biochem.* 61. 131-156.

Ayala, F.J. and Powell, J.R. 1972. Allozymes as diagnostic characters of sibling species of Drosophila. *Proc. Nat. Acad. Sci. USA.* 69(5): 1094-1096.

Barnes, W.M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc. Nat. Acad. Sci. USA* 91: 2216-2220.

Becker, J., P. Vos, M. Kuiper, F. Salamini and M. Heun. 1995. Combined mapping of AFLP and RFLP markers in barley. *Mol. Gen. Genet.* 249: 65-73.

Black, W.C. 1993. PCR with arbitrary primers: Approach with care. *Insect Mol. Biol.* 2(1): 1-6

Boreham, M.M. and Roubik, D.W. 1987. Population change and control of Africanized honeybees (Hymenoptera: Apidae) in the Panama Canal area. *Bull. Entomol. Soc. Am.* 33: 34-38.

Brown, W.M. 1985. The mitochondrial genome of animals. In Molecular Evolutionary Genetics, ed. R.J. MacIntyre. pp 95-130. Plennum Press, New York.

Carlson, D.A. and Bolten, A.B. 1983. Identification of Africanized and European honeybees using extracted hydrocarbons. *Bull. Entomol. Soc. Am.*. 30(2): 32-35.

Cavalli-Sforza, L.L., and Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. *Evolution* 32: 550-570.Cheng, S., Chang, S.Y., Gravitt, P. and Respess, R. 1994a. Long PCR. *Nature* 369: 684-685.

Cheng, S., Fockler, C., Barnes, W.M. and Higuchi, R. 1994b. Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Nat. Acad. Sci. USA* 91: 5695-5699.

Chou, Q., Russell, M., Birch, D.E., Raymond, J. and Bloch, W. 1992. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nuc. Acids Res.* 20: 1717-1723.

Collins, A.M. and Kubasek, K.J. 1982. Field test of honeybee (Hymenoptera: Apidae) colony defensive behavior. *Ann. Entomol. Soc. Am.* 75: 383-387.

Collins, A.M.; Rinderer, T.E.; Harbo, J. and Bolten, A.B. 1982. Colony defense by Africanized and European honeybees. *Science* 218: 72-74.

Contel, E.P.B., Mestriner, M.A. and Martins, E. 1977. Genetic control and developmental expression of malate dehydrogenase in *Apis mellifera*. *Biochem. Genet.* 15 (9/10): 859-876.

Crozier, R.H. 1970. On the potential for genetic variability in haplo-diploidy. *Genetica* 41: 551-556.

Crozier, Y.C., Koulianos, S., and Crozier, R.H. 1991. An improved test for Africanized honeybee mitochondrial DNA. *Experientia* 47: 968-969

Daly, H.V. 1988. Overview of the identification of the Africanized honeybees. In Africanized Honeybee and Bee Mites, pp 245-250. Ellis Horwood limited, Chechester, England.

Daly, H.V. and Balling, S.S. 1978. Identification of Africanized honeybees in the western hemisphere by discriminant analysis. *J. Kansas Entomol. Soc.* 51(4): 857-869

Daly, H.V., Hoelmer, K., Norman, P. and Allen, T. 1982. Computer-assisted measurement and identification of honeybees (Hymenoptera: Apidae). *Ann. Entom. Soc. Am.* 75: 591-594.

Danka, R.G., Hellmich, R.L. and Rindere, T.E. 1992. Nest usurpation, supersedure and colony failure contribute to Africanization of commercially managed European honeybees in Venezuela. *J. Apicult. Res.* 31(3/4): 119-123.

DeGrandi-Hoffman, G. and Watkins, J.C. 1998. Queen development time and the Africanization of European honey bees. *Am. Bee J.* 138(6): 467-469.

DeGrandi-Hoffman, G., Watkins, J.C., Collins, A.M., Loper, G.M., Martin, J.H., Arias, M.C., Sheppard, W.S. 1998. Queen developmental time as a factor in the Africanization of European honey bee (Hymenoptera : Apidae) populations. *Ann. Entomol. Soc. Am.* 91(1): 52-58.

Del Lama, M.A., Lobo, J.A., Soares, A.E.E. and Del Lama, S.N. 1990. Genetics differentiation estimated by isozymic analysis of Africanized honeybee populations from Brazil and from Central America. *Apidologie* 21, 271-280.

Eischen, F.A. 1999. Varroa and Africanized bees: an update. *Am. Bee J.* 139(8): 594-595.

Estoup, A., Solignac, M., Harry, M., and Cornuet, J.M. 1993. Characterization of $(GT)_n$ and $(CT)_n$ microsatellites in two insect species: *Apis mellifera* and *Bombus terrestris*. *Nuc. Acids. Res.* 21(6): 1427-1431.

Estoup, A., Garnery, L., Solignac, M., and Cornuet, J.M. 1995. Microsatellite variation in honey bee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics* 140: 679-695.

Excoffier, L., Smouse, P.E., and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Fletcher, D.J.C. 1973. Should European races of honey bees be introduced into South Africa? *S. Afr. Bee J.* 45: 19-25.

Fletcher, D.J.C. 1978. The African bee, *Apis mellifera adansonii* in African. *Ann. Rev. Entomol.* 23: 151-171.

Franck, P., Garnery, L., Solignac, M., and Cornuet, J.M. 1998. The Origin of west European subspecies of honeybees (*Apis mellifera*): new insights from microsatellite and mitochondrial data. *Evolution* 52(4): 1119-1134.

Garnery, L., Franck, P., Baudry, E., Vautrin, D., Cornuet, J.M., and Solignac, M. 1998a. Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A.m.iberica*). I. Mitochondrial DNA. *Genet. Sel. Evol.* 30(Suppl. 1): S31-S47.

Garnery, L., Franck, P., Baudry, E., Vautrin, D., Cornuet, J.M., and Solignac, M. 1998b. Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A.m.iberica*). II Microsatellite loci. *Genet. Sel. Evol.* 30(Suppl. 1): S49-S74.

Graur, D. 1985. Gene diversity in hymenoptera. *Evolution* 39(1): 190-199.

Guo, S.W. and Thompson, E.A. 1992. Performing an exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48: 361-372.

Hall, H.G. 1986. DNA differences found between Africanized and European honeybees. *Proc. Natl. Acad. Sci. USA* 83: 4874-4877.

Hall, H.G. 1988. Characterization of the African honeybee genotype by DNA restriction fragments. In Africanized Honeybee and Bee Mites, pp 313-324. Ellis Horwood limited.

Hall, H.G. 1990. Parental analysis of introgressive hybridization between African and European honeybees using nuclear DNA RFLPs. *Genetics* 125: 611-621.

Hall, H.G. 1992a. DNA studies reveal processes involved in the spread of New World African honeybees. *Fla. Entomol.* 75(1): 51-59.

Hall, H.G. 1992b. Further characterization of nuclear DNA RFLPs markers that distinguish African and European honeybees. *Arch. Insect. Biochem.* 19: 163-175.

Hall, H.G. 1995. Restriction fragment length polymorphism (RFLP) analyses: Distinguishing African and European honey bees. In Clapp, J.P. (ed), Molecular Methods for Species Identification, Methods in Molecular Biology, Vol. 50, Humana Press, pp. 333-366.

Hall, H.G. 1998. PCR amplification of a locus with RFLP alleles specific to African honey bees. *Biochem. Genet.* 36 (9/10): 351-361.

Hall, H.G. and Murlidharan, K. 1989. Evidences from mitochondrial DNA that African honeybees spread as continuous maternal lineages. *Nature* 339: 211-213.

Hall, H.G. and Smith, D.R. 1991. Distinguising African and European honeybee matrilines using amplified mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88: 4548-4552.

Harbo, J.R., Bolten, A.B., Rinderer, T.E. and Collins, A.M. 1981. Development periods for eggs of Africanized and European honeybees. *J. Apicult. Res.* 20(3): 156-159.

Hare, M.P., Karl, S.A., and Avise, J.C. 1996. Anonymous nuclear DNA markers in the American Oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. *Mol. Biol. Evol.* 13(2): 334-345.

Harrison, J.F. and Hall, H.G. 1993. African-European honeybee hybrids have low nonintermidate metabolic capacities. *Nature* 363: 258-260.

Her, C., and Weinshilboum, R.M. 1995. Rapid restriction mapping by use of long PCR. *Biotechniques* 19(4): 530-532.

Hill, M., H. Witsenboer, M. Zabeau, P. Vos, R. Kesseli and R. Michelmore. 1996. PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships *in Lactuca* spp. *Theor. Appl. Genet.* 93: 1202-1210.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* 6: 65-70

Hunt, G. and Page, R. 1992. Patterns of inheritance with RAPD molecular markers reveal novel types of polymorphism in the honeybee. *Theor. Appl. Genet.* 85: 15-20.

Hunt, G. and Page, R. 1995. Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. *Genetics* 139: 1371-1382.

Hunter, L.A., Jackman, J., and Sugden, E.A. 1993. Detection record of Africanized honey bees in Texas during 1990, 1991, and 1992. *Southwest. Entomol.*. 18: 79-89.

Janssen, P., R. Coopman, G. Huys, J. Swings, M. Bleeker, P. Vos, M. Zabeau and K. Kersters. 1996. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology* 142: 1881-1893.

Karl, S.A., and Avise, J.C. 1993. PCR-based assays of mendelian polymorphisms from anonymous single-copy nuclear DNA: Techniques and applications for population genetics. *Mol. Bio. Evol.* 10(2): 342-361.

Karl, S.A., Bowden, B.W. and Avise, J.C. 1992 Global population genetic structure and male mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analyses of anonymous nuclear loci. *Genetics* 131: 163-173.

Kent, R.B. 1988. The introduction and diffusion of the African honeybee in South America. *Yearbook of the Association. of Pacific Coast Geographers*. 50: 21-43.

Kerr, W.E. 1967. The history of the introduction of African bees in Brazil. *S. Afr. Bee J.* 39(2): 3-5.

Latorra, D., C.M. Stern and M.S. Schanfield. 1994. Characterization of human AFLP systems apolipoprotein B, phenylalanine hydroxylase, and D1S80. *PCR Methods Appl.* 3: 351-358.

Lavine, B.K., Carlson, D.A., Douglas, H. and Jurs, P. 1988. Taxonomy based on chemical constitution: differentiation of Africanized honeybees from European honeybees. *J. Chemometr.* 2: 29-37.

Lee, M.L. and Hall, H.G. 1996. Identification of mitochondrial DNA of Apis mellifera (Hymenoptera: Apidae) subspecies groups by multiplex allele-specific amplification with competing fluorescent-labeled primers. *Ann. Entomol. Soc. Am.* 89(1): 20-27.

Levin, M.D. 1983. Value of bee pollination to U.S. agriculture. *Bul. Entomol. Soc. Am.* 29(4): 50-51.

Lobo, J.A. and Krieger, H. 1992. Maximum likelihood estimates of gene frequencies and racial admixture in *Apis mellifera* L. (Africanized honeybees). *Heredity*. 68: 441-448.

Lobo, J.A., DelLama, M.A. and Mestriner, M.A. 1989. Population differentiation and racial admixture in the Africanized honeybee (*Apis mellifera L*). *Evolution* 43(4): 794-802.

Lockey, K.H. 1979. Insect cuticular hydrocarbons. *Com. Biochem. Physiol*. 65B: 457-462.

Majer, D., R. Mithen, B.G. Lewis, P. Vos and R.P. Oliver. 1996. The use of AFLP fingerprinting for the detection of genetic variation in fungi. *Mycol. Res.* 100: 1107-1111.

Mallet, J. 1995. A species definition for the modern synthesis. *Trends Ecol. Evol*. 10(7): 294-299.

Mc Dowell, R. 1984. The Africanized honeybee in the United States. What will happen to the U.S. beekeeping industry? USDA Agricultural report No. 519, Washington, DC.

McMichael, M. and Hall, H.G. 1996. DNA RFLPs at a highly polymorphic locus distinguish European and African subspecies of the honey bee, *Apis mellifera* L. and suggest geographical origins of New World honey bees. *Mol. Ecol.* 5:403-416.

McPherson, J.M., Eckstein, P.E., Scoles, G.J. and Gahadhar, A.A. 1993. Variability of the random amplified polymorphic DNA assay among thermal cyclers, and effects of primer and DNA concentration. *Mol. Cell. Probes.* 7: 293-299.

Mestriner, M.A. 1969. Biochemical polymorphism in bees (*Apis mellifera ligustica*). *Nature* 223: 188-189.

Mestriner, M.A. and Contel, E.P.B. 1972. The P-3 and EST loci in the honeybee *Apis mellifera*. *Genetics* 72: 733-738.

Metcalf, R.A., Marlin, J.C. and Whitt, G. 1975. Low levels of genetic heterozygosity in hymenoptera. *Nature* 257:792-795.

Michener, C.D. 1975. The Brazilian bee problem. *Ann. Rev. Entomol.* 20: 339-416.

Michener, C.D., Alfred, J., Esch, H.E., Gary, N.E., Hubell, S.P., Rothenbhuler, W.C., Smith, M.V., Towsend, G.F. and Zozaya, J.A. 1972. Final report: Committee on the African Honey Bee. Natl. Acad. Sci. USA, Washington DC.

Muller, U.G., S.E. Lipari and M.G. Milgroom. 1996. Amplified fragment length polymorphism (AFLP) fingerprinting of symbiotic fungi cultured by the fungus-growing ant *Cyphomyrmex minutus*. *Mol. Ecol.* 5: 119-122

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

Nei, M. 1987. Genetic variation within species. In Molecular and Evolutionary Genetics. pp. 176-207. Columbia University Press, New York, NY.

Nelson, E.V. 1967. History of beekeeping in the United Stated. In Beekeeping in the United Stated. Agricultural Research Service. USDA Agriculture Handbook No.335, Washington, DC.

Oertel, E. 1976. Bicentennial bee. Early records of honeybees in the Eastern United States. *Am. Bee J.* 116: 70, 71, 114, 128, 156, 157, 214, 215, 260, 261, 290.

Page, R.E. 1986. Sperm utilization in social insects. *Ann. Rev. Entomol.* 31: 297-320.

Page, R.E and Erickson, E.H. 1985. Identification and certification of Africanized honeybees. *Ann. Entomol. Soc. Am.* 78: 149-158.

Page, R.E., Erickson, E.H. and Laidlaw, H.H. 1982. A close population breeding program for honey bees. *Am. Bee J.* 122: 350-355.

Pellet, F.C. 1938. History of American Beekeeping. Collective Press, Ames Iowa.

Quezada-Euan, J.J.G., and Hinsull, S.M. 1995. Evidence of continued European morphometrics and mtDNA in feral colonies of honey bees (*Apis mellifera*) from the Yucatan peninsula, Mexico. *J. Apicult. Res.* 34: 161-166.

Quezada-Euan, J.J.G., and Medina, L.M. 1998. Hybridization between European and Africanized honeybees (*Apis mellifera* L.) in tropical Yucatan, Mexico. I. Morphometric changes in feral and managed colonies. *Apidologie*, 29 (6): 555-568.

Quezada-Euan, J.J.G., and Paxton, R.J. 1999. Rapid intergenerational changes in morphology and behaviour in colonies of Africanized and European honey bees (*Apis mellifera*) from tropical Yucatan, Mexico. *J. Apicult.Res.* 38 (1-2): 93-104

Raymond, M., and Rousset, F. 1995. Genepop (version 1.2): population genetics software for exact test and ecumenism. *J. Hered.* 86: 248-250.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43(1): 223-225

Rinderer, T.E., Collins, A., Hellmich, R.L. and Danka, R.G. 1987a. Differential drone production by Africanized and European honeybee colonies. *Apidologie*. 18(1): 61-68.

Rinderer, T.E. and Hellmich, R.L. 1988. The process of Africanization. In The African Honeybee. pp 95-117. Westview press, Boulder CO.

Rinderer, T.E., Hellmich, R.L., Danka, R.G. and Collins, A. 1985. Male reproductive parasitism: a factor in the Africanization of European honeybee populations. *Science* 228: 1119-1121.

Rinderer, T.E., Stelzer, J.A., Oldroyd, B.P., Buco, S.M. and Roubink, W.L. 1991. Hybridization between European and Africanized honeybees in the Neotropical Yucatan peninsula. *Science* 309-311.

Rinderer, T.E., Sylvester, H.A., Brown, M.A., Villa, J.D., Pesante, D. and Collins, A.M. 1986. Field and simplified technique for identyfying Africanized and European honeybees. *Apidologie* 17(1): 33-48.

Rinderer, T.E., Sylvester, H.A., Buco, M.S., Lancaster, V.A., Herbert, E.W., Collins, A.M. and Hellmich, R.L. 1987b. Improved simple techniques for identifying Africanized and European honeybees. *Apidologie* 18(2): 179-196.

Rinderer, T.E., Tucker, K.W. and Collins, A.M. 1982. Nest vavity selection by swarms of European and Africanized honeybees. *J. Apicult. Res.* 21(2): 98-103.

Ruttner, F. 1988. Biogeography and Taxonomy of Honeybees. Springer, Berlin, Germany.

Ruttner, F., Tassencourt, L., and Louveaux, J. 1978. Biometrical-statistical analysis of the geographical variability of Apis mellifera L. *Apidologie* 9**:** 363-381.

Saiki, R.K., Gelfand, D.H., Stoeffel, S., Scharf, S., Higuchi, G.T., Mullis, K.B. and Erlich, H.A. 1988. Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487-491.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular Cloning. A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanford, M.T. 1994. African bee advance in Texas stalls. *APIS* 12(7): 1

Sanford, M.T. 1999. African honey bee surprises again. *APIS* 17(5): 1

Schierwater, B. and Ender, A. 1993. Different thermostable DNA polymerases may amplify different RAPD products. *Nuc. Acids Res.* 21(19): 4647-4648.

Schiff, N.M. and Sheppard, W.S. 1993. Mitochondrial DNA evidence for the 19th century introduction of African honey bees into the United States. *Experientia* 49: 530-532.

Severson, D.W., Page, R.E. and Erickson, E.H. 1986. Closed population breeding in honey bees: a report on its practical application. *Am. Bee J.* 126**:** 93-94.

Sheppard, W.S. and Berlocher, S.H. 1985. New allozyme variability in Italian honeybees. *J. Hered* 76: 45-48

Sheppard, W.S., Rinderer, T.E., Mazzoli, J.A., Stelzer, J.A., and Shimanuki, H. 1991. Gene flow between African and European-derived honeybee populations in Argentina. *Nature*. 349:782-784.

Smith, D.R., and Brown, W.M. 1988. Polymorphism in mitochondrial DNA of European and Africanized honeybees (*Apis mellifera*). *Experientia* 44: 257-260.

Smith, D.R., and Glenn, T.C. 1995. Allozyme polymorphism in Spanish honeybees (*Apis mellifera iberica*). *J. Hered.* 86: 12-16.

Smith, D.R., Palopoli, M.F., Taylor, B.R., Garnery, L., Cornuet, J-M., Solignac, M., and Brown, W.M. 1991. Geographical overlap of two mitochondrial genomes in Spanish honebees (*Apis mellifera iberica*). *J. Hered.* 82: 96-100.

Smith, D.R., Taylor, O.R. and Brown, W.M. 1989. Neotropical Africanized honey bees have African mitochondrial DNA. *Nature* 339: 213-215.

Smith, R.K. 1990. Chemotaxonomy of Honey bees (*Apis mellifera* L.) Part 1: European and African workers. *Bee Science*. 1 (1): 23-32.

Smouse, P.E., Long, J.C., and Sokal, R.R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Syst Zool.*. 35: 627-632.

Snyder, T.P. 1974. Lack of allozyme variability in three bee species. *Evolution* 23: 687-688.

Sokal, R. and Rohlf, J. 1989. Analysis of frequencies. In Introduction to Biostatistics, 2nd ed. pp294-305 Freeman and Co, San Francisco.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments detected by gel electrophoresis. *J. Mol. Bio.* 98: 503-517.

Spivak, M. 1988. Discrimination of Africanized honeybees using behavior, cell size, morphometrics, and a newly discovered isozyme polymorphism. In Africanized Honeybee and Bee Mites. pp 313-324. Ellis Horwood Limited.

Stort, A.C. 1974. Genetic study of aggressiveness of two subspecies of *Apis mellifera* in Brazil 1. Some test to measure aggressiveness. *J. Apicult. Res.* 13(1): 33-38.

Suazo, A., McTiernan, R., and Hall, H.G. 1998. Differences between African and European honey bees (*Apis mellifera* L.) in random amplified polymorphic DNA (RAPD). *J. Hered.* 89: 32-36.

Sylvester, H.A. 1982. Electrophoretic identification of Africanized honeybees. *J. Apicult. Res.* 21(2): 93-97.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nuc. Acids. Res.* 17 (2): 6463-6471

Taylor, O.R. 1977. The past and possible future spread of Africanized honeybees in the America. *Bee World* 58: 19-30.

Taylor, O.R., Delgado, A., Brizuela, F. 1991. Rapid loss of European traits from feral neotropical African honey bee populations. *Am. Bee J.* **131**: 783-784.

Taylor, O.R. and Spivak, M. 1984. Climatic limits of tropical African bees in the Americas. *Bee World* 65: 38-47.

Thome, J., D.O. Gonzalez, S. Beebe and M.C. Duque. 1996. AFLP analysis of gene pools of a wild bean core collection. *Crop Sci.* 36: 1375-1384.

Vergara, C.D., and Perez de Leon, A. 1993. Female parasitism of European honeybees by Africanized honeybees swarms in Mexico. *J. Apicult. Res.* 32(1): 34-40.

Vos, P., R. Hodgers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nuc. Acids Res.* 23(21): 4407-4414.

Vos, P. and Kuiper, M. 1997. AFLP analysis, p115-131. *In* G. Caetano-Anolles and P. Gresshoff (Eds.), DNA markers: protocols, applications, and overviews. Wiley-Liss, New York.

Weeden, N.F. 1992. Inheritance and reliability of RAPD markers. In Applications of RAPD technology to plant breeding. pp12-17. Joint plant breeding symposia series. Minneapolis, Minnesota.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nuc. Acids Res.* 18(22): 6531-6535.

Winston, M.L., Taylor, O.R., and Otis, G.W. 1983. Some differences between temperate European and tropical African and South American honeybees. *Bee World* 64: 12-21.

BIOGRAPHICAL SKETCH

Alonso Suazo Calix was born in Tegucigalpa, Honduras, Central America, on the 20th of September 1965. He attended the Lycee Franc Hondurien in Tegucigalpa for his elementary, primary and high school education from 1972 to 1982. He continued his education at the Escuela Agricola Panamericana better known as "Zamorano" Honduras, from 1983 to 1985 where he received a degree in general agriculture, and worked for two years at "Zamorano" (1986-1987) in the beekeeping unit.

In 1988, he continued with his undergraduate studies and received a degree in entomology at the University of Florida. After graduation in 1988, he went back to continue his work in Zamorano. In January 1991, He started his Master's program in Entomology at the University of Florida and graduated in 1994.

Alonso is married to Yasmin Judith Cardoza.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

H. Glenn Hall, Chair
Associate Professor of Entomology and Nematology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gloria Moore
Professor Horticultural Science

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jane Brockmann
Professor of Zoology, department of Zoology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James Nation
Professor of Entomology and Nematology

This dissertation was submitted to the Graduate Faculty of the College of Agricultural and Life Sciences to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 2000

Dean, College of Agricultural and Life Sciences

Dean, Graduate School